

Bayesian Estimation of Multivariate Conditional Correlation GARCH models and Their Application

Chen Jun

Master's degree Programme in Bayesian Statistics and Decision Analysis
(EuroBayes)

Department of Mathematics and Statistics
University of Helsinki
November 2015



Tiedekunta/Osasto Fakultet/Sektion – Faculty		Laitos/Institution– Department	
Faculty of Science		Department of Mathematics and Statistics	
Tekijä/Författare – Author			
Chen Jun			
Työn nimi / Arbetets titel – Title			
Bayesian Estimation of Multivariate Conditional Correlation GARCH models and Their Application			
Oppiaine /Läroämne – Subject			
Master's degree Programme in Bayesian Statistics and Decision Analysis			
Työn laji/Arbetets art – Level		Aika/Datum – Month and year	Sivumäärä/ Sidoantal – Number of pages
Master's thesis		November 2015	63
Tiivistelmä/Referat – Abstract			
<p>The thesis studies three different conditional correlation Multivariate GARCH (MGARCH) models. They are the Constant Conditional Correlation (CCC-) GARCH, Dynamic Conditional Correlation (DCC-) GARCH and Asymmetric Dynamic Conditional Correlation (ADCC-) GARCH, in which the time-varying volatilities are modelled by three univariate GARCH models with the error term assumed to have a Gaussian distribution. In order to compare the performance of these models, we apply them to the volatility analysis of two stocks. Regarding model inference, we adopt a Bayesian approach and implement a Markov Chain Monte Carlo (MCMC) algorithm, Metropolis Within Gibbs (MWG), instead of the regular maximum likelihood (ML) method. Finally, the estimated models are employed to compute Value at Risk (VaR) and their performance is discussed.</p>			
Avainsanat – Nyckelord – Keywords			
CCC-GARCH, DCC-GARCH, ADCC-GARCH, MCMC			
Säilytyspaikka – Förvaringställe – Where deposited			
Kumpula science library			
Muita tietoja – Övriga uppgifter – Additional information			

Contents

1	Introduction	1
2	Univariate Models	4
2.1	Symmetric Models	4
2.1.1	ARCH Model	4
2.1.2	GARCH Model	6
2.2	Asymmetric Models	7
2.2.1	QARCH Model	7
2.2.2	GJR-ARCH Model	8
3	Multivariate Models	10
3.1	Constant Conditional Correlation Model	12
3.2	Dynamic Conditional Correlation Model	13
3.3	Asymmetric Dynamic Conditional Correlation Model	14
4	Model Estimation	16
4.1	Log likelihood function	16
4.2	Bayesian Estimation	19
4.2.1	Bayes' Theorem	19
4.2.2	Markov Chain and Monte Carlo	19

4.2.3	Main MCMC Sampling Algorithms	20
5	Empirical Analysis	24
5.1	Data	24
5.1.1	GARCH Effects	26
5.1.2	Examining Potential Asymmetry	27
5.1.3	Correlation between the returns	28
5.2	Estimation	29
5.2.1	Estimation of Static Correlation Model	31
5.2.2	Estimation of Dynamic Correlation Model	34
5.3	Model Diagnostics	35
5.4	Model Performance Comparison	36
6	Application to computing Value at Risk	39
6.1	Application in estimating VaR	39
6.1.1	In-sample test	41
6.1.2	Out-sample test	42
7	Conclusion and Future Work	45
A		47
A.1	Metropolis Hasting algorithm	47
A.2	Gibbs sampling algorithm	48
A.3	Metropolis Within Gibbs sampling algorithm	49
A.4	Sample ACF of standardised residuals	50
A.5	In-sample VaR computation	54
A.6	Out-sample VaR computation	56
A.7	Mode estimates	58
	References	61

Abstract

The thesis studies three different conditional correlation Multivariate GARCH (MGARCH) models. They are the Constant Conditional Correlation (CCC-) GARCH, Dynamic Conditional Correlation (DCC-) GARCH and Asymmetric Dynamic Conditional Correlation (ADCC-) GARCH, in which the time-varying volatilities are modelled by three univariate GARCH models with the error term assumed to have a Gaussian distribution. In order to compare the performance of these models, we apply them to the volatility analysis of two stocks. Regarding model inference, we adopt a Bayesian approach and implement a Markov Chain Monte Carlo (MCMC) algorithm, Metropolis Within Gibbs (MWG), instead of the regular maximum likelihood (ML) method. Finally, the estimated models are employed to compute Value at Risk (VaR) and their performance is discussed.

Chapter 1

Introduction

Albeit the volatility, which is a significant feature in financial time series, has been studied for several decades, the interest in it has no signs of fading. Instead, because of the increase in complexity of financial instruments, e.g. option pricing, volatility is attracting even more attention today.

A well-known property of volatility is heteroscedasticity, which is referred to the phenomenon of time-varying volatility clustering, i.e., periods of large volatility and periods of small volatility tend to appear continuously. To capture such a character in financial time series, Engle [9] introduced the Autoregressive Conditional Heteroscedasticity (ARCH) model. Whereafter, Bollerslev [3] extended the ARCH model into General Autoregressive Conditional Heteroscedasticity (GARCH) model, which achieved a huge success in modelling heteroscedasticity in financial time series. More significantly, it laid the foundation for succedent research in this topic and various extensions of the GARCH model have been developed, including, the GJR-GARCH model by Glosten, Jagannathan and Runkle [13], and the Exponential GARCH (EGARCH) model of Nelson [18].

It is straightforward to observe that volatilities move closely across assets and markets.

As we usually consider a portfolio of assets and make decisions regarding the portfolio, it is essential to study the volatilities between the assets in the portfolio simultaneously. A solution is to use univariate GARCH models to describe volatilities of each individual asset and construct the covariance matrix between the assets. This is the basic idea used to develop a Multivariate GARCH (MGARCH) model where main challenge is to make the model sufficiently flexible and yet parsimonious.

Of the multivariate GARCH models proposed the VEC-GARCH model of Bollerslev, Engle, and Wooldridge [5] and its restricted version the BEKK model of Engle and Kroner [10] are flexible but not parsimonious. In this thesis we therefore consider multivariate GARCH model based on the conditional correlations.

In 1990, Bollerslev [4] introduced a multivariate GARCH model, referred as the Constant Conditional Correlation (CCC-) GARCH model, in which univariate GARCH models related to one another with a constant correlation matrix. In his Dynamic Conditional Correlation (DCC-) GARCH models, Engle [8] relaxed the assumption of a constant correlation by allowing for time varying correlations. A further extension is the Asymmetric Dynamic Conditional Correlation (ADCC-) GARCH model proposed by Cappiello, Engle and Sheppard [6] with the motivation of allowing for asymmetric impacts on the correlations caused by good news and bad news.

Generally, the classical approach of estimating the parameters of GARCH models is the method of ML. However, in this thesis a Bayesian approach: Markov Chain Monte Carlo (MCMC) method is applied. Ardia and Hoogerheide [1] outline the MCMC method for GARCH models. Of various MCMC algorithms applied in the estimation of GARCH models, Tetsuya Takaishi [21] employed the Hamilton Monte Carlo algorithm and Ardia [2] used Metropolis Hasting scheme in GARCH(1,1) with Student- t innovations. They demonstrated the MCMC method can be very useful in estimating GARCH models.

This thesis consists of the following chapters: In Chapter 2 and Chapter 3, the uni-

variate GARCH models and multivariate GARCH models considered in the thesis are introduced. In Chapter 4, Bayesian estimation is discussed in the context of multivariate GARCH models. In Chapter 5, CCC-GARCH, DCC-GARCH and ADCC-GARCH models are built for two stocks, General Electric and American Express Company, and applied in Chapter 6 to estimate VaR. Finally, conclusions are provided in Chapter 7.

Chapter 2

Univariate Models

2.1 Symmetric Models

2.1.1 ARCH Model

The traditional assumption that the one-period forecast error variance is constant across all the time points in a series does not reflect the existence of time varying variance observed in stock returns. To allow for this feature, Engle [9] introduced the autoregressive conditional heteroskedasticity (ARCH) model. The ARCH model assumes that conditional variances given the past information are non-constant described by a function of previous squared returns. The definition of the q th order ARCH model, $\text{ARCH}(q)$ for short, is given in the following definition:

Definition 2.1.1. The ARCH(q) model is defined by the equations

$$r_t = \mu_t + y_t \quad (2.1)$$

$$y_t = \sqrt{h_t} z_t, \quad (2.2)$$

$$h_t = \omega + \sum_{i=1}^q \alpha_i y_{t-i}^2, \quad (2.3)$$

where

- r_t is the considered return at time t ;
- μ_t is the expected return at time t ;
- y_t is the mean corrected return at time t ;
- z_t is a sequence of independent and identical distributed (i.i.d.) with zero mean and unit variance, such that z_t is independent of $(r_{t-1}, r_{t-2}, \dots)$;
- h_t is the conditional variance of the return at time t given the past returns $(r_{t-1}, r_{t-2}, \dots)$;
- $\omega > 0$ and $\alpha_i \geq 0$ are parameters used to specify the conditional variance h_t with the inequalities guaranteeing that h_t is positive;

In Equation (2.1), μ_t could be time varying and modelled by some time series model, e.g., an autoregressive model. However, because our main object is to study the volatility, we assume that μ is constant and equal to the mean of the series. In Equation (2.2), z_t is transformed into y_t by multiplying with $\sqrt{h_t}$, and from the assumption on z_t made it follows that h_t defined in (2.3) is the conditional variance of the returns. Thus the model describes the heteroskedasticity observed in reality.

From the definition of the ARCH model, it is explicit that for $\alpha_i > 0$, a large value of a squared past return y_{t-i}^2 results in a large conditional variance h_t and the probability that the mean corrected return y_t takes a large absolute value increases. This corresponds to

the clustering phenomenon observed in financial time series. After the introduction of the ARCH model, its usefulness in describing time varying volatility was established in numerous applications.

2.1.2 GARCH Model

Although the ARCH model made extraordinary progress in the analysis of time varying volatility, it has some limitations. Tsay [23] showed that it often requires many lags to describe the volatility adequately, i.e., q in Equation (2.3) should be large. In 1986, Tim Bollerslev [3] introduced an extension, the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model, in which the conditional variance is not only explained by past squared returns but also by past conditional variances, thereby allowing for a more flexibility.

Definition 2.1.2. The definition of the GARCH(p, q) is defined by the equations:

$$r_t = \mu_t + y_t, \quad (2.4)$$

$$y_t = \sqrt{h_t} z_t, \quad (2.5)$$

$$h_t = \omega + \sum_{i=1}^p \alpha_i y_{t-i}^2 + \sum_{i=1}^q \beta_i h_{t-i}, \quad (2.6)$$

where, r_t, μ_t, y_t and z_t are as in the definition of the ARCH model (See Definition 2.1.1).

In the GARCH model, h_t is explained by both squared past returns y_{t-i}^2 and past conditional variances h_{t-i} . To ensure h_t is greater than zero, $\omega > 0$ and $\alpha_i, \beta_i \geq 0$ are assumed and if $\alpha_i = 0$ for all i , then $\beta_i = 0$ for all i must hold. Empirical experience has shown that less free parameters are needed in the GARCH model than in the ARCH model to describe the volatility of the same asset return, and consequently the GARCH model has been used more widely in practice than the ARCH model [23].

2.2 Asymmetric Models

In financial markets, the leverage effect is used to refer to the unequal effect caused by bad news and good news. To allow for the leverage effect, extensions of the GARCH model have been developed, including the EGARCH model of Nelson [18], the NGARCH model of Engle and Ng [11] and the Threshold GARCH model of Zakoian [24]. In the thesis, we selected two asymmetric GARCH models, the QGARCH model of Sentana [20] and the GJR-GARCH model of Glosten, Jagannathan and Runkle [13] are considered.

2.2.1 QARCH Model

The QGARCH model is short for the Quadratic GARCH model proposed by Sentana [20]. Unlike the conventional GARCH model, the QGARCH model allows the conditional variance to depend not only on squared returns but also on the returns. A formal definition of the QGARCH model is as follows

Definition 2.2.1. The definition of the QGARCH(p, q) is defined by the equations:

$$r_t = \mu_t + y_t \quad (2.7)$$

$$y_t = \sqrt{h_t} z_t \quad (2.8)$$

$$h_t = \omega + \sum_{i=1}^p \alpha_{ii} y_{t-i}^2 + \sum_{i=1}^p \phi_i y_{t-i} + 2 \sum_{i=1}^p \sum_{j=i+1}^p \alpha_{ij} y_{t-i} y_{t-j} + \sum_{i=1}^q \beta_i h_{t-i} \quad (2.9)$$

where, r_t, μ_t, y_t and z_t are as in the definition of the ARCH model (See Definition 2.1.1).

As seen in Equation (2.6) and Equation (2.9), the QGARCH model contains the extra part $\sum_{i=1}^p \phi_i y_{t-i} + 2 \sum_{i=1}^p \sum_{j=i+1}^p \alpha_{ij} y_{t-i} y_{t-j}$ not appearing in the GARCH model.

In order to derive the condition for the positive variance, we rewrite Equation (2.9) in

vector form as

$$h_t = \omega + \mathbf{Y}'_{t-1} \mathbf{A} \mathbf{Y}_{t-1} + \boldsymbol{\phi}' \mathbf{Y}_{t-1} + \sum_{i=1}^q \beta_i h_{t-i}. \quad (2.10)$$

$$= (\mathbf{Y}'_{t-1}, 1) \begin{bmatrix} \mathbf{A} & \frac{\boldsymbol{\phi}}{2} \\ \frac{\boldsymbol{\phi}'}{2} & \omega \end{bmatrix} \begin{pmatrix} \mathbf{Y}_{t-1} \\ 1 \end{pmatrix} + \sum_{i=1}^q \beta_i h_{t-i}, \quad (2.11)$$

where $\boldsymbol{\phi}$ is a $p \times 1$ vector with the components ϕ_1, \dots, ϕ_p and \mathbf{A} is a symmetric $p \times p$ matrix with elements $\alpha_{i,j} = \alpha_{j,i}$. To satisfy the positivity assumption of the conditional variance h_t , we require $\omega > 0$, $\beta_i \geq 0$, and that the matrices $\begin{bmatrix} \mathbf{A} & \frac{\boldsymbol{\phi}}{2} \\ \frac{\boldsymbol{\phi}'}{2} & \omega \end{bmatrix}$ and \mathbf{A} are positive semidefinite. If \mathbf{A}^+ denotes the Moore-Penrose inverse of \mathbf{A} , the non-negativity of the is guaranteed by the inequality $\omega - \boldsymbol{\phi}' \mathbf{A}^+ \boldsymbol{\phi} / 4 \geq 0$ [20].

With the purpose to have a parsimonious model, we always treat the matrix \mathbf{A} as a diagonal matrix in which case the conditional variance in Equation (2.9) simplies to

$$h_t = \omega + \sum_{i=1}^p \alpha_{ii} y_{t-i}^2 + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{i=1}^q \beta_i h_{t-i} \quad (2.12)$$

To see how the QGARCH model works, consider Equation(2.12), where the coefficient ϕ_i shows the difference to the GARCH model. As usual assume that bad news brings negative mean corrected returns y_{t-i}^- and good news brings positive returns y_{t-i}^+ . Then if ϕ_i is negative, the inequality $\phi_i y_{t-i}^- > \phi_i y_{t-i}^+$ is satisfied, and it follows from Equation (2.12) that the negative return y_{t-i}^- has a greater impact on the volatility than a positive returns y_{t-i}^+ .

2.2.2 GJR-ARCH Model

Another model widely used to describe the leverage effect is the GJR-GARCH model, it is named after Glosten, Jagannathan, and Runkle [13] and defined as follows.

Definition 2.2.2. The GJR-GARCH(p, q) model is defined by the equations

$$r_t = \mu_t + y_t \quad (2.13)$$

$$y_t = \sqrt{h_t} z_t, \quad (2.14)$$

$$h_t = \omega + \sum_{i=1}^p \alpha_i y_{t-i}^2 + \sum_{i=1}^p \gamma_i I_i[y_{t-i} < 0] y_{t-i}^2 + \sum_{i=1}^q \beta_i h_{t-i} \quad (2.15)$$

where, r_t, μ_t, y_t and z_t are as in the definition of the ARCH model (See Definition 2.1.1) and

$$I_i[y_{t-i} < 0] = \begin{cases} 1 & \text{if } y_{t-i} < 0 \\ 0 & \text{if } y_{t-i} \geq 0 \end{cases} \quad (2.16)$$

From the definition of GJR-GARCH model, it is seen a negative mean corrected return y_{t-i} caused by bad news will aggrandise the conditional variance h_t by $\gamma_i y_{t-i}^2$.

Compared to the QGARCH model, the GJR-GARCH model has the advantage that the restrictions on the parameters needed to guarantee positive conditional variance are relative simpler, only $\omega > 0$ and that all other parameters are non-negative will suffice.

Chapter 3

Multivariate Models

Generally, financial returns move simultaneously over time across different assets or markets and their comovement are dependent on their covariance structure. Consequently, in empirical analyses, especially in financial risk management, studying the volatilities with multivariate models is more relevant than working with univariate models separately.

Before presenting specific multivariate GARCH models, we describe the general structure used in the definition.

Definition 3.0.1. The general structure of models for multivariate conditional heteroskedasticity is given by the equations

$$\mathbf{r}_t = \boldsymbol{\mu}_t + \mathbf{y}_t \tag{3.1}$$

$$\mathbf{y}_t = \mathbf{H}_t^{1/2} \mathbf{z}_t \tag{3.2}$$

where

- \mathbf{r}_t is the $n \times 1$ vector of returns at time t ;
- $\boldsymbol{\mu}_t$ is the $n \times 1$ vector of expected returns at time t ;

- \mathbf{y}_t is the $n \times 1$ vector of mean corrected returns at time t ;
- \mathbf{z}_t is a $n \times 1$ sequence of i.i.d random vectors with zero mean and identity covariance matrix and such that \mathbf{z}_t is independent of past returns $(\mathbf{r}_1, \mathbf{r}_2, \dots)$;
- \mathbf{H}_t is the $n \times n$ conditional covariance matrix of the returns at time t given the past returns $(\mathbf{r}_1, \mathbf{r}_2, \dots)$.

The Equation (3.1) and Equation (3.2) have obvious counterparts in the univariate models discussed in previous chapters. For example, the vector $\boldsymbol{\mu}_t$ contains the expected returns and \mathbf{H}_t is a multivariate extension of the conditional covariance. To complete the specification of the model, the conditional covariance matrix \mathbf{H}_t should be specified.

As mentioned in the Introduction, examples of the specifications proposed in literature, include the VEC model [5] and the BEKK model [10]. But in the thesis, multivariate conditional correlation models will be considered.

In the multivariate conditional correlation volatility models, the conditional covariance matrix \mathbf{H}_t is specified in a two-step way. Firstly, one specifies the conditional variances separately for each returns by using a univariate GARCH model. For instance, the conditional variances could follow a GARCH model or a QGARCH model if allowance of asymmetric effects is desired. Whereafter, the conditional variances specified in the first step are used to build the conditional covariance matrix \mathbf{H}_t . So the following sections we discuss the multivariate conditional correlation models, namely the Constant Conditional Correlation model, the Dynamic Conditional Correlation model and the Asymmetric Conditional Correlation model.

3.1 Constant Conditional Correlation Model

The Constant Conditional Correlation (CCC) GARCH model was first put forward by Bollerslev [4]. It is the simplest model among the three models to be discussed. The key assumption is that the conditional correlations between the elements of \mathbf{y}_t are time invariant.

Definition 3.1.1. Given the Definition 3.0.1, the specification of \mathbf{H}_t in CCC-GARCH model is given by equations

$$\mathbf{H}_t = \mathbf{D}_t^{1/2} \mathbf{R} \mathbf{D}_t^{1/2} \quad (3.3)$$

$$\mathbf{D}_t = \text{diag}(h_{11,t}, \dots, h_{nn,t}) \quad (3.4)$$

where

1. \mathbf{D}_t is a diagonal matrix with positive diagonal entries that are the conditional variances specified by univariate GARCH models;
2. \mathbf{R} is a positive definite correlation matrix, i.e.,

$$\mathbf{R} = \begin{pmatrix} 1 & \rho_{12} & \dots & \rho_{1n} \\ \rho_{21} & 1 & \dots & \rho_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1} & \rho_{n2} & \dots & 1 \end{pmatrix} \quad (3.5)$$

In Definition 3.0.1, \mathbf{H}_t is required to be positive definite. In the CCC-GARCH model, this condition is guaranteed by the requirement \mathbf{R} and \mathbf{D}_t are positive definite, this particularly means that all conditional variances must be positive so that the constraints for positive conditional variance should be satisfied for each univariate GARCH model. From Equation (3.3), we have

$$h_{ij,t} = \rho_{ij} \cdot \sqrt{h_{ii,t} h_{jj,t}}, \quad i \neq j, \quad (3.6)$$

and when \mathbf{R} equals the identity matrix \mathbf{I} so that $\rho_{ij} = 0$ when $i \neq j$, we get the special case where all assets are independent with conditional variances modeled by univariate GARCH models.

3.2 Dynamic Conditional Correlation Model

Compared to other multivariate GARCH models, the CCC-GARCH model reduces the number of free parameters. However, the assumption of constant conditional correlation is restrictive in most empirical applications. This leads to Engle's [8] the Dynamic Conditional Correlation (DCC)-GARCH model, in which the correlation matrix \mathbf{R} is time variant.

Definition 3.2.1. Given Definition 3.0.1, the specification of \mathbf{H}_t in the DCC-GARCH model is given by the equations

$$\mathbf{H}_t = \mathbf{D}_t^{1/2} \mathbf{R}_t \mathbf{D}_t^{1/2} \quad (3.7)$$

$$\mathbf{D}_t = \text{diag}(h_{11,t}, \dots, h_{nn,t}) \quad (3.8)$$

$$\mathbf{R}_t = \text{diag}(q_{11,t}^{-1/2}, \dots, q_{nn,t}^{-1/2}) \mathbf{Q}_t \text{diag}(q_{11,t}^{-1/2}, \dots, q_{nn,t}^{-1/2}) \quad (3.9)$$

$$\mathbf{Q}_t = (1 - a - b) \bar{\mathbf{Q}} + a \boldsymbol{\epsilon}_{t-1} \boldsymbol{\epsilon}_{t-1}' + b \mathbf{Q}_{t-1} \quad (3.10)$$

where

1. \mathbf{D}_t is as in the CCC-GARCH model (see Definition 3.1.1);
2. \mathbf{R}_t is the time-varying correlation matrix

$$\mathbf{R}_t = \begin{pmatrix} 1 & \rho_{12,t} & \dots & \rho_{1n,t} \\ \rho_{21,t} & 1 & \dots & \rho_{2n,t} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1,t} & \rho_{n2,t} & \dots & 1 \end{pmatrix} \quad (3.11)$$

and $q_{ii,t}$ in Equation (3.9) is the i th diagonal element of the symmetric matrix \mathbf{Q}_t with typical element $q_{ij,t}$

3. $\boldsymbol{\epsilon}_t$ is a standardised residual vector with component $\epsilon_{i,t} = y_{i,t} / \sqrt{h_{ii,t}}$;
4. $\bar{\mathbf{Q}}$ is a positive definite parameter matrix with unit diagonal elements.

According to the Equation (3.9) and Equation (3.10), the correlation matrix \mathbf{R}_t is constructed through the matrix \mathbf{Q}_t . In addition to a well defined conditional variances, the scalar parameters a, b need to be non-negative and satisfy $a + b < 1$ to guarantee \mathbf{Q}_t is positive definite and consequently \mathbf{R}_t is positive. Positive constraints on \mathbf{D}_t and \mathbf{R}_t ensure the positive definite of \mathbf{H}_t .

3.3 Asymmetric Dynamic Conditional Correlation Model

Similar to asymmetries in volatility, asymmetric effect also exists in the conditional correlations. This feature can be captured by the Asymmetric DCC-GARCH model of Cappiello, Engle and Sheppard [6].

Definition 3.3.1. Given Definition 3.0.1, the specification of \mathbf{H}_t in the ADCC-GARCH model is given by the equations

$$\mathbf{H}_t = \mathbf{D}_t^{1/2} \mathbf{R}_t \mathbf{D}_t^{1/2} \quad (3.12)$$

$$\mathbf{D}_t = \text{diag}(h_{11,t}, \dots, h_{nn,t}) \quad (3.13)$$

$$\mathbf{R}_t = \text{diag}(q_{11,t}^{-1/2}, \dots, q_{nn,t}^{-1/2}) \mathbf{Q}_t \text{diag}(q_{11,t}^{-1/2}, \dots, q_{nn,t}^{-1/2}) \quad (3.14)$$

$$\mathbf{Q}_t = (\bar{\mathbf{Q}} - A' \bar{\mathbf{Q}} A - B' \bar{\mathbf{Q}} B - G' \bar{\mathbf{N}} G) + A' \boldsymbol{\epsilon}_{t-1} \boldsymbol{\epsilon}_{t-1}' A + B' \mathbf{Q}_{t-1} B + G' \mathbf{n}_{t-1} \mathbf{n}_{t-1}' G \quad (3.15)$$

where the quantities in Equation (3.12), Equation (3.13) and Equation (3.14) are as the DCC-GARCH model (see Definition 3.2.1). In Equation (3.15),

1. $\mathbf{n}_t = \mathbf{I}[\boldsymbol{\epsilon}_t < 0] \circ \boldsymbol{\epsilon}_t$ (\circ is the Hadamard product) with the i th component of the $n \times 1$ vector $\mathbf{I}[\boldsymbol{\epsilon}_t < 0]$ defined as

$$I_i[\epsilon_{i,t} < 0] = \begin{cases} 1 & \text{if } \epsilon_{i,t} < 0 \\ 0 & \text{if } \epsilon_{i,t} \geq 0 \end{cases} \quad (3.16)$$

2. A , B and G are diagonal parameter matrices.
3. $\bar{\mathbf{N}}$ is a positive semidefinite parameter matrix.

In order to guarantee that the conditional covariance matrix \mathbf{H}_t is positive definite, the constraints on the matrix \mathbf{D}_t and \mathbf{R}_t should be the same as in DCC-GARCH model. In the ADCC-GARCH model, as \mathbf{R}_t needs to be positive definite, matrix \mathbf{Q}_t needs to be positive definite, and therefore the following condition should hold:

$$\bar{\mathbf{Q}} - AA' \circ \bar{\mathbf{Q}} - BB' \circ \bar{\mathbf{Q}} - GG' \circ \bar{\mathbf{N}} \quad (3.17)$$

is positive definite.

With the purpose to reduce the dimension of the parameter vector but still keep the flexibility of the model, we adopt a special case and replace the coefficient matrix A , B , G by scalars \sqrt{a} , \sqrt{b} and \sqrt{g} . Then the Equation (3.15) can be rewritten as

$$\mathbf{Q}_t = (\bar{\mathbf{Q}} - a\bar{\mathbf{Q}} - b\bar{\mathbf{Q}} - g\bar{\mathbf{N}}) + a\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}_{t-1}' + b\mathbf{Q}_{t-1} + g\mathbf{n}_{t-1}\mathbf{n}_{t-1}' \quad (3.18)$$

and the condition for positive definition of \mathbf{Q}_t can be rewritten as: $a + b + \delta g < 1$, where δ is the maximum eigenvalue $[\bar{\mathbf{Q}}^{-1/2}\bar{\mathbf{N}}\bar{\mathbf{Q}}^{-1/2}]$ [6].

Chapter 4

Model Estimation

In this chapter, we describe the scheme used to estimate parameters of the models introduced in the preceding chapters. A essential assumption is that the standard error \mathbf{z}_t follows a standardised multivariate Gaussian distribution.

4.1 Log likelihood function

When the standardised error \mathbf{z}_t is from a multivariate Gaussian distribution with $E[\mathbf{z}_t] = 0$ and $\text{Var}[\mathbf{z}_t] = \mathbf{I}$, the joint distribution of an independent sample $\mathbf{z}_1, \dots, \mathbf{z}_T$ is:

$$f(\mathbf{z}_1, \dots, \mathbf{z}_T) = \prod_{i=1}^T \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2} \mathbf{z}_i' \mathbf{z}_i\right) \quad (4.1)$$

According to Definition 3.0.1, we assume that the expected return $\boldsymbol{\mu}_t$ equals a constant so that $\mathbf{r}_t = \boldsymbol{\mu} + \mathbf{y}_t$. To simplify notation, we assume in this chapter that $\boldsymbol{\mu} = 0$ and use the notation \mathbf{y}_t for the vector of returns.

Using the properties of the multivariate Gaussian distribution, when $\mathbf{y}_t = \mathbf{H}_t^{1/2} \mathbf{z}_t$, we have

$\text{Var}[\mathbf{y}_t] = \mathbf{H}_t$ and the joint distribution of $\mathbf{y}_1, \dots, \mathbf{y}_T$ is:

$$f(\mathbf{y}_1, \dots, \mathbf{y}_T) = \prod_{i=1}^T \frac{1}{(2\pi)^{n/2} |\mathbf{H}_t|^{1/2}} \exp\left(-\frac{1}{2} \mathbf{y}_t' \mathbf{H}_t^{-1} \mathbf{y}_t\right), \quad (4.2)$$

where \mathbf{H}_t is determined by the models discussed in the thesis.

In independent univariate GARCH models (see the end of section 3.1), the correlation matrix \mathbf{R}_t equals to an identity matrix \mathbf{I} and $\mathbf{H}_t = \mathbf{D}_t = \text{diag}(h_{11,t}, \dots, h_{nn,t})$. Thus, from (5.2) we get the log likelihood function:

$$\begin{aligned} \log(L(\mathbf{Y}|\boldsymbol{\varphi})) &= -\frac{1}{2} \sum_{i=1}^T (n \log(2\pi) + \sum_{j=1}^n (\log(h_{jj,t}) + \frac{y_j^2}{h_{jj,t}})) \\ &= \sum_{j=1}^n [-\frac{1}{2} \sum_{i=1}^T (\log(2\pi) + \log(h_{jj,t}) + \frac{y_j^2}{h_{jj,t}})] \end{aligned} \quad (4.3)$$

where $\mathbf{Y} = (\mathbf{y}_1, \dots, \mathbf{y}_T)'$ is the $T \times n$ data vector and the parameter vector $\boldsymbol{\varphi}$ contains the parameters in the considered univariate GARCH models $h_{11,t}, \dots, h_{nn,t}$. For instance, $h_{ii,t}$ can follow conventional GARCH models in which case $\boldsymbol{\varphi}$ contains the parameters ω_i, α_i , and β_i ($i = 1, \dots, n$). The above equation shows that in this case the log likelihood function of all the assets is the sum of log likelihood functions of each single asset.

As an extension of the independent univariate model but still in the family of static correlation models, we have the CCC-GARCH model whose log likelihood function is obtained from (4.2) by choosing $\mathbf{D}_t^{1/2} \mathbf{R} \mathbf{D}_t^{1/2}$ with \mathbf{D}_t and \mathbf{R} as in Definition 4.0.2 (\mathbf{D}_t is a diagonal matrix as specified above). The log likelihood function follows the form

$$\log(L(\mathbf{Y}|\boldsymbol{\varphi}, \boldsymbol{\rho})) = -\frac{1}{2} \sum_{i=1}^T (n \log(2\pi) + \log |\mathbf{D}_t^{1/2} \mathbf{R} \mathbf{D}_t^{1/2}| + \mathbf{y}_t' (\mathbf{D}_t^{1/2} \mathbf{R} \mathbf{D}_t^{1/2})^{-1} \mathbf{y}_t) \quad (4.4)$$

where $\boldsymbol{\varphi}$ is as in (4.3) and $\boldsymbol{\rho}$ contains the elements in the correlation matrix \mathbf{R} .

The likelihood function of a DCC-GARCH or ADCC-GARCH model is difficult to handle

in practice. Following Engle [8], we therefore divide the log likelihood function into two parts.

$$\begin{aligned}\log L(\mathbf{Y}|\boldsymbol{\theta}) &= -\frac{1}{2} \sum_{i=1}^T (n \log(2\pi) + 2 \log |\mathbf{D}_t| + \log |\mathbf{R}_t| + \boldsymbol{\epsilon}_t' \mathbf{R}_t^{-1} \boldsymbol{\epsilon}_t) \\ &= -\frac{1}{2} \sum_{i=1}^T (n \log(2\pi) + 2 \log |\mathbf{D}_t| + \mathbf{y}_t' (\mathbf{D}_t \mathbf{D}_t)^{-1} \mathbf{y}_t - \boldsymbol{\epsilon}_t' \boldsymbol{\epsilon}_t + \log |\mathbf{R}_t| + \boldsymbol{\epsilon}_t' \mathbf{R}_t^{-1} \boldsymbol{\epsilon}_t)\end{aligned}\tag{4.5}$$

where, $\boldsymbol{\epsilon}_t = \mathbf{D}_t^{-1/2} \mathbf{y}_t$ is the vector of standardised returns, \mathbf{D}_t is as in Equation (3.4) and \mathbf{R}_t is defined in the Definition 3.2.1 or Definition 3.3.1. Furthermore $\boldsymbol{\theta} = (\boldsymbol{\varphi}, \boldsymbol{\psi})$ is the parameter vector that contains the parameters $\boldsymbol{\varphi}$ in the considered univariate GARCH model whereas parameters $\boldsymbol{\psi}$ contains the parameters in the DCC-GARCH models or ADCC-GARCH models in Definition 3.2.1 or Definition 3.3.1. We notice that in the Equation (4.5), the correlation matrix only appears in some terms, so that we can write:

$$\log L(\mathbf{Y}|\boldsymbol{\theta}) = \log L_v(\mathbf{Y}|\boldsymbol{\varphi}) + \log L_c(\mathbf{Y}|\boldsymbol{\psi})\tag{4.6}$$

where

$$\log L_v(\mathbf{Y}|\boldsymbol{\varphi}) = -\frac{1}{2} \sum_{i=1}^T (n \log(2\pi) + 2 \log |\mathbf{D}_t| + \mathbf{y}_t' (\mathbf{D}_t^{1/2} \mathbf{D}_t^{1/2})^{-1} \mathbf{y}_t)\tag{4.7}$$

equals to the log likelihood function for univariate model in (4.3) and the correlation part

$$\log L_c(\mathbf{Y}|\boldsymbol{\psi}) = -\frac{1}{2} \sum_{i=1}^T (\log |\mathbf{R}_t| + \boldsymbol{\epsilon}_t' \mathbf{R}_t^{-1} \boldsymbol{\epsilon}_t - \boldsymbol{\epsilon}_t' \boldsymbol{\epsilon}_t)\tag{4.8}$$

We divide the log-likelihood function of the dynamic conditional correlation models into two parts and will introduce the two-step method to estimate parameters in Chapter 5.

4.2 Bayesian Estimation

A classic method to estimate GARCH models is Maximum Likelihood Estimation (MLE) method. Different from the MLE, the rapid development in Bayesian estimation techniques based on Markov Chain and Monte Carlo (MCMC) method offers an option to estimate GARCH models.

4.2.1 Bayes' Theorem

Assume $\boldsymbol{\theta} = (\theta_1, \dots, \theta_n)$ is the parameter vector with a prior of density function $\pi(\boldsymbol{\theta})$ and $\mathbf{Y} = (\mathbf{y}_1, \dots, \mathbf{y}_T)'$ is the $T \times n$ data vector. Based on the Bayes' rule, the posterior density $\pi(\boldsymbol{\theta}|\mathbf{Y})$ is given by:

$$\pi(\boldsymbol{\theta}|\mathbf{Y}) \propto L(\mathbf{Y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}), \quad (4.9)$$

where $L(\mathbf{Y}|\boldsymbol{\theta})$ is the likelihood function. Generally, the posterior $\pi(\boldsymbol{\theta}|\mathbf{Y})$ is unnormalised and the normalization can be done by $\pi(\boldsymbol{\theta}|\mathbf{Y})/Z$ with $Z = \int \pi(\boldsymbol{\theta}|\mathbf{Y})d\boldsymbol{\theta}$.

In cases, when we have no beliefs on the parameters, we can use independent prior uninformative to all unknown parameters. Then let the prior $\pi(\boldsymbol{\theta})$ be proportional to a constant, so that according to the Bayes' theorem,

$$\pi(\boldsymbol{\theta}|\mathbf{y}) \propto L(\mathbf{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}) \propto L(\mathbf{y}|\boldsymbol{\theta}), \quad (4.10)$$

Consequently, the posterior distribution of the parameter vector $\boldsymbol{\theta}$ is proportional to the likelihood function when independent uninformative priors are selected.

4.2.2 Markov Chain and Monte Carlo

The main idea behind the MCMC method is to set up an ergodic Markov Chain which has the posterior distribution as its stationary distribution. Using the Markov Chain, one

then simulates a large sample

$$\boldsymbol{\theta}^{(0)}, \boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(v+k)}, \quad (4.11)$$

and deletes the first $v - 1$ samples with $v + k$ large to obtain

$$\boldsymbol{\theta}^{(v)}, \boldsymbol{\theta}^{(v+1)}, \dots, \boldsymbol{\theta}^{(v+k)} \quad (4.12)$$

which will be used for estimation. The reason of deleting the first $v - 1$ samples is that, they vary heavily and their distribution is far from the stationary distribution, considering them in the estimate will bias the result. Using $\boldsymbol{\theta}^{(v)}, \boldsymbol{\theta}^{(v+1)}, \dots, \boldsymbol{\theta}^{(v+k)}$ with $v+k$ "large", we can calculate the expectation of the posterior distribution, quantiles and other statistics. Theoretically speaking, the expectation of the posterior, which will be used to estimate the parameters in our model, can be calculated

$$\hat{\boldsymbol{\theta}} = E(\boldsymbol{\theta}|\mathbf{y}) = \frac{1}{Z} \int \boldsymbol{\theta} \pi(\boldsymbol{\theta}|\mathbf{y}) d\boldsymbol{\theta} \quad (4.13)$$

However, when the posterior is high dimensional or otherwise complicated, it may be very difficult to analyse it in closed form and obtain the normalising constant Z , and hence the estimate. As this is also the case in our model, we use the samples from the MCMC method to obtain a numerical approximation for the expectation of the posterior distribution. By the case of large numbers, so that the estimate used is

$$\hat{\boldsymbol{\theta}} = \frac{1}{k} \sum_{i=1}^k \boldsymbol{\theta}^{(v+i)} \quad (4.14)$$

with k is large.

4.2.3 Main MCMC Sampling Algorithms

In this section, I will introduce two main MCMC algorithms and a combination of them applied in the thesis.

Metropolis-Hastings algorithm

The Metropolis-Hastings (M-H) algorithm is a popular MCMC algorithms used to obtain a sequence of random samples from a target distribution, typically a posterior distribution, for which direct sampling is difficult. The algorithm is first proposed by Nicholas Metropolis, ect. [17] and extended by W. K. Hastings [14]. To show how the algorithm works, consider generating a sample from the target distribution $\pi(\boldsymbol{\theta}|\mathbf{Y})$. In the M-H algorithm, we need a proposed distribution $q(\cdot|\boldsymbol{\theta})$ defined on the parameter space of $\boldsymbol{\theta}$. Assuming the current state $\boldsymbol{\theta} = \boldsymbol{\theta}^{(i)}$, we propose a value $\boldsymbol{\theta}'$ for the next state from the proposed distribution $q(\boldsymbol{\theta}'|\boldsymbol{\theta})$ and accept it with a probability

$$p = \min\left\{1, \frac{\pi(\boldsymbol{\theta}'|\mathbf{Y}) \cdot q(\boldsymbol{\theta}|\boldsymbol{\theta}')}{\pi(\boldsymbol{\theta}|\mathbf{Y}) \cdot q(\boldsymbol{\theta}'|\boldsymbol{\theta})}\right\} \quad (4.15)$$

To check if $\boldsymbol{\theta}'$ is accepted, we sample a random variable u from a standard uniform distribution, and we accept $\boldsymbol{\theta}'$ if $u < p$. Otherwise, we reject $\boldsymbol{\theta}'$. The next current state $\boldsymbol{\theta} = \boldsymbol{\theta}^{(i+1)} = \boldsymbol{\theta}'$ only if $\boldsymbol{\theta}'$ is accepted. If $\boldsymbol{\theta}'$ is rejected, the previous current state remains, that is, $\boldsymbol{\theta} = \boldsymbol{\theta}^{(i)}$. We repeat these steps until a sufficient number of samples as obtained.

A special case of the M-H algorithm is a random walk algorithm, in which the proposed distribution has the symmetric property $q(\boldsymbol{\theta}|\boldsymbol{\theta}') = q(\boldsymbol{\theta}'|\boldsymbol{\theta})$. For example, in a normal random walk, we draw samples $\mathbf{W} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}$ is a covariance matrix, and propose the value $\boldsymbol{\theta}' = \boldsymbol{\theta}^{(i)} + \sigma\mathbf{W}$ where σ is a user chosen size parameter. As $q(\boldsymbol{\theta}|\boldsymbol{\theta}') = q(\boldsymbol{\theta}'|\boldsymbol{\theta})$, the acceptance probability p in (4.15) simplifies to:

$$p = \min\left\{1, \frac{\pi(\boldsymbol{\theta}'|\mathbf{Y})}{\pi(\boldsymbol{\theta}|\mathbf{Y})}\right\} \quad (4.16)$$

A code for the M-H algorithm is provided in Appendix A.1.

Gibbs sampling

Another widely used MCMC algorithm is Gibbs sampling named after Josiah Willard Gibbs but proposed by Stuart and Geman [12]. Before the introduction of Gibbs sampling, we introduce the concept of full conditionals. For a parameter vector $\boldsymbol{\theta} = (\theta_1, \dots, \theta_n)$, the full conditional of θ_i in the posterior $\pi(\boldsymbol{\theta}|\mathbf{Y})$ is defined as:

$$\pi(\theta_i|\theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_n, \mathbf{Y}) = \pi(\theta_i|\boldsymbol{\theta}_{-i}, \mathbf{Y}), \quad (4.17)$$

where $\boldsymbol{\theta}_{-i}$ is the parameter vector obtained from $\boldsymbol{\theta}$ by removing the element θ_i . Interpreting the joint distribution $\pi(\theta_1, \dots, \theta_i, \dots, \theta_n, \mathbf{Y})$ as a function of θ_i shows that

$$\pi(\theta_i|\theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_n, \mathbf{Y}) \propto \pi(\boldsymbol{\theta}, \mathbf{Y}) \quad (4.18)$$

A critical condition for Gibbs sampling is that, for all $i = 1, \dots, n$, sampling from the conditional distributions in (4.17) is possible. Once this condition is satisfied, we can apply Gibbs sampling.

The idea behind Gibbs sampling is to simulate successively each component θ_i from its conditional distribution. We make use of the most recent value $\boldsymbol{\theta}_{-i}$ and update θ_i with its value sampled from the conditional distribution $\pi(\theta_i|\boldsymbol{\theta}_{-i}, \mathbf{Y})$. After the update of θ_i , i.e., the i th element in the parameter vector $\boldsymbol{\theta}$, we update the next element $i + 1$ from its full conditional distribution $\pi(\theta_{i+1}|\boldsymbol{\theta}_{-(i+1)}, \mathbf{Y})$. We repeat the sampling consecutively until we have enough samples for all the elements in the parameter vector $\boldsymbol{\theta}$ to make the approximation of true value. Code in Appendix A.2 shows how the Gibbs Sampling works.

Metropolis Within Gibbs algorithm

The Metropolis Hasting and Gibbs sampling have their advantages. The M-H algorithm works even if the posterior is complicated and sampling from full conditional distribution

is impossible. Regarding the Gibbs sampling algorithm, it simplifies high dimensional problems by successively generating from different subsets of $\boldsymbol{\theta}$. However, even if we apply both methods, we cannot ignore the disadvantages. In a high dimension model, the M-H algorithm is inefficient because the increase in the dimension decreases the acceptance rate, whereas the Gibbs sampling is infeasible when full conditionals, i.e., the conditional distribution in (4.17) are unknown. In order to find an efficient technique, we employ the Metropolis-Within-Gibbs algorithm in the thesis to assimilate advantages in each technique.

Metropolis-within-Gibbs (MWG) is a hybrid MCMC algorithm that combines M-H and Gibbs sampling [22]. The idea of MWG is to execute a Metropolis step when a random sample from its full conditionals is failed in Gibbs sampling, i.e. when (4.17) are unknown, instead of updating θ_i by sampling from the its posterior full conditionals, we then accept a new state θ'_i by a Metropolis step with the probability p in Equation (4.15). The MWG algorithm is more efficient in high dimensions cases than the M-H algorithm because a Gibbs step handle the low acceptance rate problem caused by high dimension. On the other hand, in the case when the full conditionals are unknown and it is impossible to implement a Gibbs sampling step, a M-H step can be used. Therefore, the MWG algorithm is a good option in the estimation of different GARCH models considered in the thesis. Code in Appendix A.3 shows the Metropolis-Within-Gibbs with a normal random walk.

Chapter 5

Empirical Analysis

In this chapter, the models and methods discussed in previous chapters are applied to analyze stock market data.

5.1 Data

The data selected for the analysis are the daily log-returns of two stocks from the New York Stock Exchange, American Express Company (AXP) and General Electric Company (GE). The data are obtained from Yahoo Finance. To guarantee a reasonably large sample, the date covers five years from January 1st, 2005, to December 9th, 2010. As the two time series should have the same length, we eliminate the days without a daily price in both stocks. After preliminary processing, we have 1495 daily observations in total.

Since the main goal is to analyze volatility, the series are first transformed to log-return as $100 \cdot \ln(p_t/p_{t-1})$, where p_t is the closed price. Thereafter the sample means of the log-returns are subtracted from the series, so that the data finally be employed in the

empirical analysis are the mean corrected log-returns. If not specified, we will use returns to refer to mean corrected log returns in the following analysis. Figure 5.1 depicts the returns. The first impression is that during some periods, both returns have fluctuated heavily.

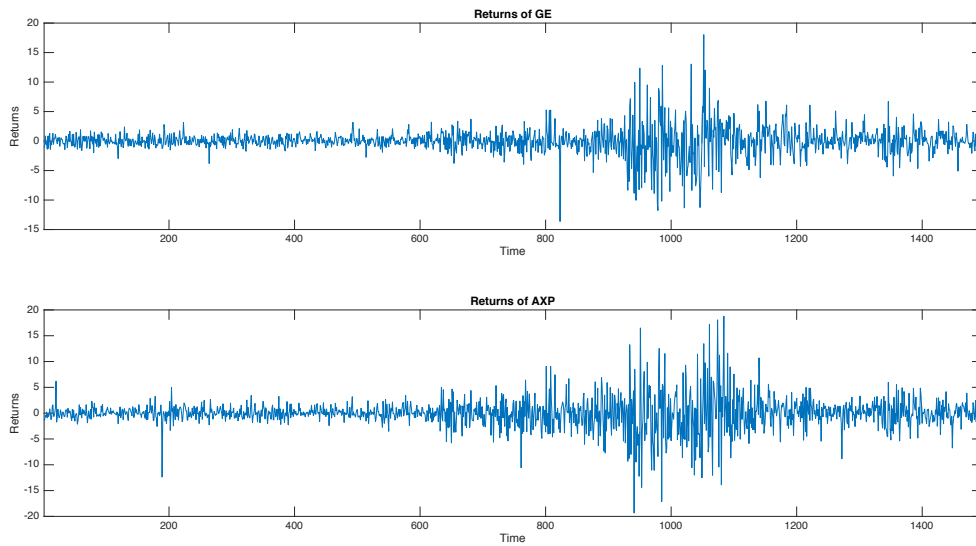


Figure 5.1: The GE and AXP returns

Figure 5.2 shows the histogram of both equities and the estimated normal distribution density function. We can conclude that the distribution of the returns does not follow a Gaussian distribution because there are large gaps between the fitted density function and the histogram. Table 5.1 presents descriptive statistics of the returns. Due to the mean correction, the means of both returns are zero and therefore not shown in the table. For both returns, the skewness is positive which suggests that the right tail of the distribution is fatter than the left tail. In addition, the kurtosis of both returns are greater than three which similarity to the histogram points to leptokurtic distributions with fat tails. Finally, the p -values of the Jarque-Bera test are almost zero indicating non-normality of both returns.

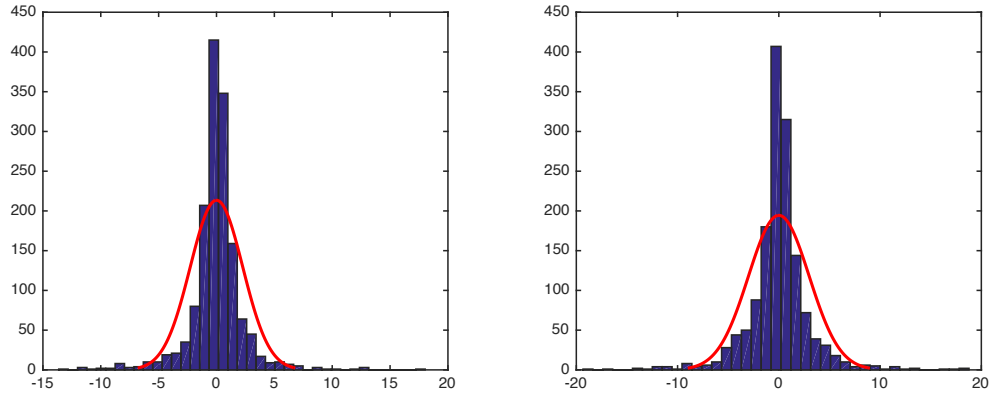


Figure 5.2: Histogram of the GE and AXP returns and the estimated density function of normal distribution

Descriptive Statistics	GE Returns	AXP Returns
St.Dev	2.27	3.00
Median	0.02	0.01
Min	-13.63	-19.34
Max	18.04	18.78
Skewness	0.02	0.08
Kurtosis	9.77	8.03
Jarque-Bera	$p\text{-value} = 0.001$	$p\text{-value} = 0.001$

Table 5.1: Descriptive Statistics of the GE and AXP Returns

5.1.1 GARCH Effects

Figure 5.1 shows the typical volatility clustering observed in stock returns. In absolute value, large returns tend to be followed by large returns and small returns tend to be followed by small returns. Using the sample autocorrelation functions (ACF) of the series and their squares, Figure 5.3 shows that another typical feature of stocks returns holds true for the GE and AXP series. In the ACF plots, the vertical lines show the sample

autocorrelations at different lags and the horizontal lines show the 0.05 critical bound. Sample autocorrelations within the critical bounds supports lack of correlation for both series. Figure 5.3 shows lack of autocorrelation in returns but strong autocorrelation in squared returns. This is a feature that can be captured by a GARCH model.

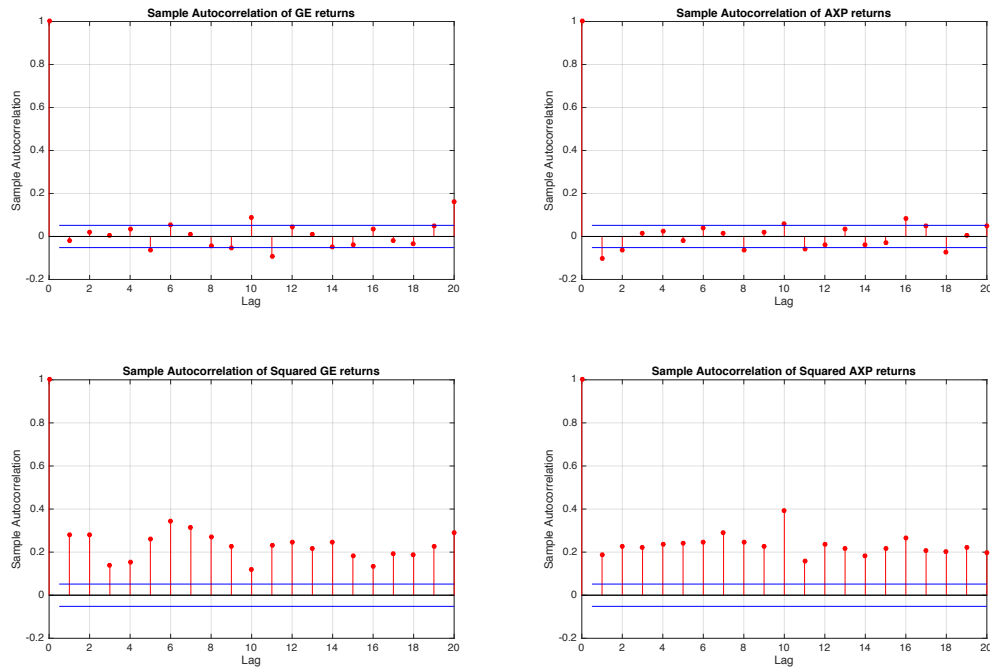


Figure 5.3: Autocorrelation functions of the GE and AXP returns: sample ACF of GE returns (above left), sample ACF of squared GE returns (below left), sample ACF of AXP returns (above right) and sample ACF of squared AXP returns (below right).

5.1.2 Examining Potential Asymmetry

Before applying GARCH models, it is of interest to examine potential leverage effects, the existence of which would suggest the use of an asymmetric GARCH model in analysis. We

use the simple idea of Zivot [25] and examine the sample correlation between the squared return y_t^2 and the lagged return y_{t-1} . A non-zero value of this suggests potential leverage effects exist in the time series.

Series	Correlation
GE	-0.0181
AXP	-0.1032

Table 5.2: Correlation between y_t^2 and y_{t-1}

The estimates obtained for the GE returns and AXP returns are -0.0181 and -0.1032 respectively. Thus, no strong indication about the leverage effect is obtained, which can be testified by the estimated parameters in the asymmetric GARCH models in the estimation part.

5.1.3 Correlation between the returns

By comparing the time series graphs of the two returns, we find that they exhibit similarities. For example, between observations 1 and 600, both volatility are small whereas between observations 800 and 1200 both of them are much larger. This supports the idea that the returns are dependent. Figure 5.4, (a), (b) and (c) shows scatterplots of the returns. The plot in (a) shows that the two returns are contemporaneously correlated while they are uncorrelated one of them is lagged with on day. This demonstrates the rationality to apply the multivariate models introduced in the previous part chapters.

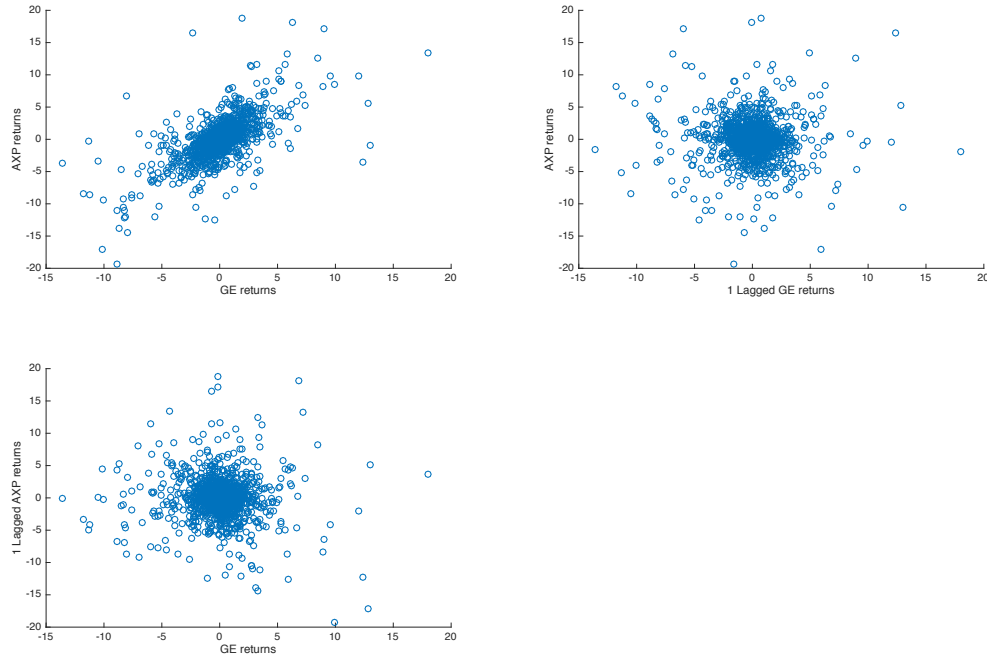


Figure 5.4: Scatter plot of GE and AXP: (a). The scatter plot of the GE and AXP returns (above left). (b). The scatter plot of the 1-lagged return of GE and the AXP return (above right). (c). The scatter plot of the 1-lagged return of AXP and the GE return (below left).

5.2 Estimation

In this section, the models introduced in Chapter 3 are applied to the observed data. Initially, we employ parsimonious GARCH models, i.e., GARCH(1,1), QGARCH(1,1) and GJR-GARCH(1,1). After that we consider multivariate models starting with the static models, i.e., the CCC-GARCH model and its special case, independent univariate GARCH models. Thereafter, we focus on the dynamic correlation models, i.e., DCC-GARCH models and ADCC-GARCH models.

For parameter estimation, we use the Bayesian estimation technique with the Metropolis-

within-Gibbs algorithm discussed in section 4.2.3. As we have no information on the prior, we follow the approach mentioned in 4.2.1 and adopt independent uninformative priors and implementing the algorithm on the likelihood function.

Meanwhile, we also need to impose parameter restrictions on the sample procedure to guarantee the employed models are stationary and the conditional variances are positive as well as the conditional correlation matrices are positive definite. Table 5.3, Table 5.4 and Table 5.5 summarize the conditions for each in the each models. To understand

Model	Parameter restrictions
GARCH(1,1)	$\alpha + \beta < 1$
QGARCH(1,1)	$\alpha + \beta < 1$
GJR-GARCH(1,1)	$\alpha + \frac{\gamma}{2} + \beta < 1$

Table 5.3: Restrictions for stationarity [3], [20], [16]

Model	Parameter restrictions
GARCH(1,1)	$\omega, \alpha, \beta > 0$
QGARCH(1,1)	$\omega, \alpha, \beta > 0, 4\alpha\omega - \phi^2 \geq 0$
GJR-GARCH(1,1)	$\omega, \alpha, \beta > 0$

Table 5.4: Restrictions for positive conditional variance (See Chapter 2)

CCC	$-1 < \rho_{ij} < 1$
DCC	$a + b \leq 1, a, b > 0$
ADCC	$a + b + \delta g < 1, a, b, g > 0$

Table 5.5: Restrictions for positive definite conditional correlation matrix (See Chapter 3)

and use these conditions, we select the CCC-GARCH(1,1) model as an example. Firstly, to satisfy the stationary condition and positive conditional variance, it is essential to ensure $\alpha + \beta < 1$ and $\omega, \alpha, \beta > 0$ for both of the two GARCH(1,1) models. Secondly, to ensure that the correlation matrix is positive definite, we need $-1 < \rho_{12} < 1$.

Finally, we generate sample of size 5,000 for the parameter vector $\boldsymbol{\theta}$ with burn-in the first 1000 samples. As mentioned in the section 4.2.2, we adopt the expectation of the posterior samples as an estimate of $\boldsymbol{\theta}$. However, when we shall compare the performance of different models with Bayesian Information Criterion (BIC) [19] we need the maximum likelihood estimate and for this purpose we use to the mode of the posterior samples which should provide a close approximation for the ML estimation [7].

5.2.1 Estimation of Static Correlation Model

After the preliminary analysis, now we have a multivariate return series $\mathbf{y}_t = (y_{1,t}, y_{2,t})'$, $t = 1, \dots, T$, in which $y_{1,t}$ is the return of GE and $y_{2,t}$ is the return of AXP. Table 5.6 and Table 5.7 presents parameter estimates for the independent univariate GARCH model and the CCC-GARCH model respectively (Numbers in parentheses below estimates are standard errors obtained as standard deviations from the MWG samples). The estimates of β_i are around 0.9 in all models implying strong conditional heteroskedasticity in both returns. Furthermore, the negative estimates of ϕ_i in the QGARCH model and positive estimates of γ_i in the GJR-GARCH model suggest to leverage effects. The leverage effects are not very strong in some cases, however, and this is particularly the case for the first component of the CCC-GJR-GARCH model in Table 5.7. In the CCC-GARCH model, the positive estimated of the correlation coefficient ρ implies that the two stocks are related with one of the returns typically following the other one in the same direction.

GE&AXP Daily returns Jan 1st2005-December 9th 2010 (1,496 obs.)			
$\mathbf{y}_t = \mathbf{H}_t^{1/2} \mathbf{z}_t$			
$h_{11,t} = \omega_1 + \alpha_1 y_{1,t-1}^2 + \phi_1 y_{1,t-1} + \gamma_1 I[y_{1,t-1} < 0] + \beta_1 h_{11,t-1}$			
$h_{22,t} = \omega_2 + \alpha_2 y_{2,t-1}^2 + \phi_2 y_{2,t-1} + \gamma_2 I[y_{2,t-1} < 0] + \beta_2 h_{22,t-1}$			
$h_{ij,t} = 0, \text{ when } i \neq j$			
Parameter	GARCH(1,1)	QGARCH(1,1)	GJR-GARCH(1,1)
ω_1	0.0187 (0.0080)	0.0258 (0.0088)	0.0220 (0.0086)
α_1	0.0871 (0.0170)	0.0911 (0.0167)	0.0699 (0.0170)
ϕ_1		-0.0703 (0.0240)	
γ_1			0.0426 (0.0216)
β_1	0.9126 (0.0158)	0.9070 (0.0154)	0.9084 (0.0164)
ω_2	0.0353 (0.0132)	0.0465 (0.0146)	0.0421 (0.0129)
α_2	0.0965 (0.0159)	0.1006 (0.0160)	0.0547
ϕ_2		-0.1022 (0.0335)	
γ_2			0.0924 (0.0337)
β_2	0.9049 (0.0140)	0.8996 (0.0138)	0.9013 (0.0134)
Maximum Log likelihood	-5978.532	-5967.076	-5971.468
BIC	12000.92	11992.63	12001.42

Table 5.6: Parameter estimates for univariate GARCH(1,1) models.

GE&AXP Daily returns Jan 1st2005-December 9th 2010 (1,496 obs.)			
$\mathbf{y}_t = \mathbf{H}_t^{1/2} \mathbf{z}_t$			
$h_{11,t} = \omega_1 + \alpha_1 y_{1,t-1}^2 + \phi_1 y_{1,t-1} + \gamma_1 I[y_{1,t-1} < 0] + \beta_1 h_{11,t-1}$			
$h_{22,t} = \omega_2 + \alpha_2 y_{2,t-1}^2 + \phi_2 y_{2,t-1} + \gamma_2 I[y_{2,t-1} < 0] + \beta_2 h_{22,t-1}$			
$h_{ij,t} = \rho, \text{ when } i \neq j$			
Parameter	GARCH(1,1)	QGARCH(1,1)	GJR-GARCH(1,1)
ω_1	0.0230 (0.0078)	0.0282 (0.0084)	0.0267 (0.0086)
α_1	0.0718 (0.0140)	0.0760 (0.0135)	0.0688 (0.0153)
ϕ_1		-0.0485 (0.0242)	
γ_1			0.0189 (0.0134)
β_1	0.9225 (0.0140)	0.9173 (0.0135)	0.9157 (0.0154)
ω_2	0.0520 (0.0143)	0.0600 (0.0157)	0.0558 (0.0150)
α_2	0.0817 (0.0135)	0.0832 (0.0132)	0.0601 (0.0165)
ϕ_2		-0.0877 (0.0358)	
γ_2			0.0459 (0.0248)
β_2	0.9112 (0.0129)	0.9087 (0.0128)	0.9096 (0.0134)
ρ	0.5857 (0.0198)	0.5824 (0.0198)	0.5828 (0.0198)
Maximum Log likelihood	-5672.387	-5666.931	-5670.907
BIC	11395.94	11399.65	11407.6

Table 5.7: Parameter estimation for CCC-GARCH models.

Below the estimates, the maximum value of the log likelihood function and BIC are given for each model (see the discussion at the end of the preceding section), these will be discussed when comparing all estimated models in section 5.4.

5.2.2 Estimation of Dynamic Correlation Model

To estimate the parameters of DCC-GARCH and ADCC-GARCH models, we employ the two-step method proposed by Engle [8]. The first step is to estimate a univariate GARCH model and use the resulting estimates in the second step to estimate the remaining parameters. Specifically, we first implement the MWG algorithm on the univariate GARCH model, then the first component on the right hand side of (4.6), $\log L_v(\mathbf{Y}|\boldsymbol{\varphi})$ becomes a fixed constant after the first step. Once we have estimated the parameters for the univariate GARCH models, Engle [8] suggests that we can use $\frac{1}{n} \sum_{t=1}^T \boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t'$ to estimate $\bar{\mathbf{Q}}$ and $\frac{1}{n} \sum_{t=1}^T \mathbf{n}_t \mathbf{n}_t'$ to estimate $\bar{\mathbf{N}}$ (see Definition 3.2.1 and Definition 3.3.1). Therefore, it suffices to run the MWG algorithm on the second component in (4.6), $\log L_c(\mathbf{Y}|\boldsymbol{\psi})$. The first step estimates we use are those given in Table 5.6. Those estimates are used in computing the the log-likelihood function (4.7), which is applied to DCC-GARCH models and ADCC-GARCH models estimation. Table (5.8) and Table (5.9) present the estimation results obtained in this way (Numbers in parentheses below estimates are standard errors obtained as standard deviations from the MWG samples).

In both dynamic conditional correlation models, the estimates of b are around 0.8 while the estimates of a are close to 0.02. This shows that the Q_t matrix is relative stable as it depends much on its previous state and fluctuation caused by the standardised residual is small. In the ADCC-GARCH(1,1) model, the estimates of g are exist and around 0.02, which supports the existence of asymmetric effects caused by negative and positive returns, but these effects are not strong.

GE&AXP Daily returns Jan 1st 2005-December 9th 2010 (1,496 obs.)			
$Q_t = (1 - a - b)\bar{Q} + a\epsilon_{t-1}\epsilon'_{t-1} + bQ_{t-1}$			
Parameter	GARCH(1,1)	QGARCH(1,1)	GJR-GARCH(1,1)
a	0.0288 (0.0142)	0.0268 (0.0138)	0.0281 (0.0143)
b	0.8220 (0.1427)	0.8374 (0.1328)	0.8281 (0.1204)
\bar{Q}	$\begin{bmatrix} 0.9785 & 0.5661 \\ 0.5661 & 0.9859 \end{bmatrix}$	$\begin{bmatrix} 0.9782 & 0.5614 \\ 0.5614 & 0.9848 \end{bmatrix}$	$\begin{bmatrix} 0.9781 & 0.5586 \\ 0.5586 & 0.9822 \end{bmatrix}$
Maximum Log likelihood	-5673.275	-5668.506	-5674.137
BIC	11412.34	11417.43	11428.69

Table 5.8: Parameter estimates for DCC-GARCH models based on the univariate models in Table 5.6

5.3 Model Diagnostics

After having estimated the parameters of a model, we need to check the adequacy of the model. Regarding this, we show in Appendix A.4 autocorrelation functions of empirical counterpart of the standardized residuals obtained as

$$\hat{\mathbf{z}}_t = \mathbf{y}_t \hat{\mathbf{H}}_t^{-1/2} \quad (5.1)$$

where \mathbf{y}_t is the vector of returns and $\hat{\mathbf{H}}_t$ is the conditional covariance matrix based on the estimated models. In this case autocorrelations of the squared residuals are of interest and a model is deemed adequate if autocorrelations of squared standardised residuals are within the critical bound similar to those in Figure 5.3. In the ACF plot in Appendix A.4, they show no autocorrelation in the squared returns, i.e, no inadequacies in the estimated models.

GE&AXP Daily returns Jan 1st 2005-December 9th 2010 (1,496 obs.)			
$Q_t = (\bar{Q} - a\bar{Q} - b\bar{Q} - g\bar{N}) + a\epsilon_{t-1}\epsilon'_{t-1} + bQ_{t-1} + g\mathbf{n}_{t-1}\mathbf{n}_{t-1}$			
Parameter	GARCH(1,1)	QGARCH(1,1)	GJR-GARCH(1,1)
a	0.0207 (0.0137)	0.0172 (0.0129)	0.0238 (0.0140)
b	0.7959 (0.2043)	0.8609 (0.1464)	0.7725 (0.1681)
g	0.0159 (0.0125)	0.0189 (0.0130)	0.0115 (0.0100)
\hat{Q}	$\begin{bmatrix} 0.9785 & 0.5661 \\ 0.5661 & 0.9859 \end{bmatrix}$	$\begin{bmatrix} 0.9782 & 0.5614 \\ 0.5614 & 0.9848 \end{bmatrix}$	$\begin{bmatrix} 0.9781 & 0.5586 \\ 0.5586 & 0.9822 \end{bmatrix}$
\hat{N}	$\begin{bmatrix} 0.4848 & 0.3192 \\ 0.3192 & 0.5184 \end{bmatrix}$	$\begin{bmatrix} 0.4864 & 0.3192 \\ 0.3192 & 0.5184 \end{bmatrix}$	$\begin{bmatrix} 0.4663 & 0.1747 \\ 0.1747 & 0.5418 \end{bmatrix}$
Maximum Log likelihood	-5674.869	-5668.896	-5675.508
BIC	11444.78	11447.45	11460.67

Table 5.9: Parameter estimates for ADCC-GARCH models based on the univariate models in Table 5.6

5.4 Model Performance Comparison

In this section, we compare the performance of the estimated models using two criteria: the maximum value of the log likelihood function and the Bayesian Information Criterion (BIC).

By the definition, the maximum value of the likelihood function \hat{L} measures how well a model fits the data in comparison with alternative models, a high value of the likelihood function indicates goodness of fit. However, the maximized value of likelihood function does not allow for the complexity of the model caused by increasing the num-

ber of parameters, that is overfitting. The second criterion adopted to investigate the model performance is BIC, where a penalty term is used to guard against using too many parameters. The BIC, developed by Gideon E. Schwarz [19], is defined by:

$$\text{BIC} = -2 \cdot \ln \hat{L} + k \cdot \ln(n) \quad (5.2)$$

where

- \hat{L} is the maximised value of the likelihood function of the model.
- k is the number of free parameters estimated.
- n is the number of observations data.

Models with small values of BIC are recommended. In (5.2), $k \cdot \ln(n)$ is the introduced penalty term which penalises using an excessive number of parameters in the model. Achieving a high value of the likelihood function by increasing the number of parameters may not result in a favourable model in terms of BIC because the value of the penalty term is also increased.

Referring to the value of likelihood function and BIC computed by the mode of the posterior samples, comparing the static correlation models in terms of the value of the maximized likelihood function shows that the asymmetric GARCH(1,1) models are preferable to the symmetric GARCH(1,1) models. In addition, we can conclude that the CCC-GARCH models are better at capturing the series' volatility than the corresponding independent univariate GARCH models. Regarding the BIC values of the univariate models, we notice that, in spite of having a higher likelihood value, the GJR-GARCH(1,1) model is not preferred by BIC in comparison to the GARCH(1,1) model which, however is outperformed by the QGARCH(1,1) model. As a group, the CCC-GARCH models still beat the univariate models according to BIC. Comparing the three CCC-GARCH models, the CCC-GARCH(1,1) turns out to be the best, followed by the CCC-QGARCH(1,1) model and CCC-GJR-GARCH(1,1) model.

Of the dynamic correlation models, the DCC-GARCH models beat the ADCC-GARCH models according to likelihood value as well as according to BIC. In each group, models with the QGARCH(1,1) model used to describe the conditional variance are the best according to the likelihood value but the BIC suggests the more parsimonious DCC-GARCH(1,1) or ADCC-GARCH(1,1) model. From the definition of the DCC-GARCH and ADCC-GARCH models, we know that DCC-GARCH models are nested in ADCC-GARCH models so that according to maximised values of the likelihood function, ADCC-GARCH models should be preferred. A possible reason why this is not the case is that in the two-step estimation, $\bar{\mathbf{Q}}$ and $\bar{\mathbf{N}}$ are not maximum likelihood estimates and bias exists.

Finally, we discuss the performance across the static correlation models and dynamic correlation models. If we check the performance according to the maximum likelihood value, the CCC-QGARCH(1,1) model is selected. As the CCC-GARCH model is a special case of DCC-GARCH model, the DCC-GARCH MODEL would be preferred by the maximum likelihood value theoretically, but comparing all the CCC-GARCH models and DCC-GARCH models, we find that the maximum likelihood value chooses the CCC-GARCH model. This is another evidence to show that the two-step method is biased. As to the other criterion, BIC, it still recommends the CCC-GARCH(1,1) model. An interpretation is that correlation exists in the returns but a constant correlation specification is sufficient to capture it. In addition, the leverage effect exists but it is not strong and BIC insists that THE symmetric GARCH model is sufficient to describe the volatility in the two returns series.

Chapter 6

Application to computing Value at Risk

In this section, we compare the performance of the estimated multivariate GARCH models in computing the in sample Value at Risk and forecasting out sample Value at Risk.

6.1 Application in estimating VaR

VaR, which is short for Value at Risk, is one of the most popular techniques used in risk management to estimate the worst expected loss at a given probability level. We define the VaR of a portfolio:

Definition 6.1.1. Let r_t be a daily return at day t and $\eta \in (0, 1]$. Then the VaR is a real number VaR_η such that:

$$P_{t-1}(r_t > \text{VaR}_\eta) = 1 - \eta \tag{6.1}$$

where $P_{t-1}(\cdot)$ denotes the conditional probability given the available information at day $t - 1$.

If we have a single asset and the returns follows the definition $r_t = \mu + \sqrt{h_t}z_t$ with h_t and z_t independent and $z_t \sim N(0, 1)$, we can compute the VaR as follows

$$\text{VaR}_\eta = \mu + \sqrt{h_t}z_{t,\eta} \quad (6.2)$$

where $z_{t,\eta}$ is the η -quantile of the distribution of z_t .

The definition of VaR is usually given for a single index or a single asset. But in practice, it is often used for a portfolio of several assets. This can be done as follows: suppose we have n different risky assets in a portfolio with a weight vector $\mathbf{w} = (w_1, \dots, w_n)'$ and assume these assets' returns follow Definition 3.0.1 so that $\mathbf{r}_t = \boldsymbol{\mu} + \mathbf{H}_t^{1/2}\mathbf{z}_t$, where \mathbf{H}_t is the conditional covariance matrix, $\mathbf{z}_t \sim N(\mathbf{0}, \mathbf{I})$ and \mathbf{z}_t and \mathbf{H}_t independent. Consequently, we can derive the distribution of the portfolio by using the properties of multivariate normal distribution. Specifically, the portfolio $\mathbf{w}'\mathbf{r}_t$ follows a univariate normal distribution with mean $\mathbf{w}'\boldsymbol{\mu}$ and conditional variance $\mathbf{w}'\mathbf{H}_t\mathbf{w}$ and Equation 6.2 can be rewritten as:

$$\text{VaR}_\eta = \mathbf{w}'\boldsymbol{\mu} + z_{t,\eta}\sqrt{\mathbf{w}'\mathbf{H}_t\mathbf{w}} \quad (6.3)$$

Thus, the VaR also works as a risk measure of multivariate returns.

After the calculation of VaR, there are many tests which can be used to compare the performance of comparing models in terms of VaR. We use the test of Kupiec [15], which compare the relative number of violations in VaR with the chosen probability level η . Here a violation means that the daily return is less than the VaR value of the day. If the number of VaR's computed is N and the number of violations is x , it is reasonable to treat x as a Boernoulli distributed random variable with parameter η . The idea of the Kupiec test is to compare the observed violation rate $\hat{\eta} = x/N$ with the chosen probability level η . The considered model is in doubt if the observed failure rate $\hat{\eta}$ is significantly different

from its theoretical counterpart η . Kupiec built the test on the likelihood ratio:

$$\text{LR} = 2 \ln \frac{\hat{\eta}^x (1 - \hat{\eta})^{N-x}}{\eta^x (1 - \eta)^{N-x}} \quad (6.4)$$

In the case of a correctly specified model, test statistic LR is asymptotically chi-square distributed with one degree of freedom and large values of the test statistic are critical.

Now we are ready to discuss how we apply the Kupiec test. We select $\eta = 0.95$ and assume that both stocks GE and AXP are held in the portfolio with equal 50% shares. As we subtract the mean of returns in the preliminary analysis, $\boldsymbol{\mu} = (0, \dots, 0)'$ and the returns are assumed to follow a conditional normal distribution with mean zero and covariance matrix is \mathbf{H}_t , hence Equation (6.3) can be simplified as:

$$\text{VaR}_{0.95} = 1.65 \sqrt{\mathbf{w}' \mathbf{H}_t \mathbf{w}} \quad (6.5)$$

where the theoretical conditional covariance matrix \mathbf{H}_t is replaced by an estimate obtained from the estimated model.

In the following two sections, we first apply the Kupiec test within the sample period and then out of the sample period. We use the 5% significant level so that the critical value of the chi-square distribution with one degree freedom is 3.84. Thus, a value of the test statistic greater than 3.84 casts doubt on the adequacy of the model.

6.1.1 In-sample test

In the in-sample test, we will compute the conditional covariance matrix \mathbf{H}_t and compute the $\text{VaR}_{0.95}$ according to Equation (6.5). The figures in Appendix A.5 show time series the $\text{VaR}_{0.95}$ computed with different models and the composed returns of both stocks within the sample data.

Table 6.1 shows the results of in-sample Kupiec test.

Models	Relative number of violations	LR Value
GARCH(1,1)	0.088	36.8333
QGARCH(1,1)	0.086	34.4582
GJR-GARCH(1,1)	0.087	35.6373
DCC-GARCH(1,1)	0.046	0.4693
DCC-QGARCH(1,1)	0.046	0.4693
DCC-GJR-GARCH(1,1)	0.045	0.6513
ADCC-GARCH(1,1)	0.053	0.4693
ADCC-QGARCH(1,1)	0.053	0.4693
ADCC-GJR-GARCH(1,1)	0.045	0.8642
CCC-GARCH(1,1)	0.047	0.1960
CCC-QGARCH(1,1)	0.045	0.6513
CCC-GJR-GARCH(1,1)	0.044	1.1083

Table 6.1: Results of the in-sample Kupiec test with critical value of the LR test statistic is 3.84

The results show that the univariate models perform much worse than the multivariate models. The difference in the LR value among the univariate models is rather small and the same is true for the multivariate models. The results indicates that allowing for the correlation between the two returns is advantageous.

6.1.2 Out-sample test

In this section, we implement out-of-sample Kupiec test by using the returns of GE and AXP from 9th December 2010 to 9th December 2014, with totally 1006 daily log returns. In computing the conditional covariance matrix \mathbf{H}_t to estimate the VaR, the same esti-

Models	Relative number of violations	LR Value
GARCH(1,1)	0.060	2.2502
QGARCH(1,1)	0.059	1.5043
GJR-GARCH(1,1)	0.063	3.1359
DCC-GARCH(1,1)	0.033	7.0925
DCC-QGARCH(1,1)	0.032	8.0030
DCC-GJR-GARCH(1,1)	0.033	7.0925
ADCC-GARCH(1,1)	0.034	6.2446
ADCC-QGARCH(1,1)	0.031	8.7981
ADCC-GJR-GARCH(1,1)	0.033	7.0925
CCC-GARCH(1,1)	0.029	11.1301
CCC-QGARCH(1,1)	0.028	12.3115
CCC-GJR-GARCH(1,1)	0.028	12.3115

Table 6.2: Results of the out-sample Kupiec test with critical value of the LR test statistic is 3.84

mates from the previous in-sample data are used for all models without update. Figures in Appendix A.5 present the VaRs based on one day ahead forecasts and returns, whereas Table 6.2 displays the test results. The out-of-sample results in Table 6.2 are completely contrary to the in-sample results in Table 6.1 in that the out-of-sample Kupiec test clearly rejects all multivariate models. A possible reason is that without updating estimates, the multivariate GARCH models lose accuracy in the forecasting of the conditional covariance matrix \mathbf{H}_t when the prediction period is too long. To examine this assumption, we do the Kupiec test for the first 300 out-of-sample and have the results in Table 6.3.

Models	Relative number of violations	LR Value
GARCH(1,1)	0.070	2.2590
QGARCH(1,1)	0.067	1.5955
GJR-GARCH(1,1)	0.070	2.2590
DCC-GARCH(1,1)	0.033	1.9779
DCC-QGARCH(1,1)	0.033	1.9779
DCC-GJR-GARCH(1,1)	0.033	1.9779
ADCC-GARCH(1,1)	0.037	1.2325
ADCC-QGARCH(1,1)	0.033	1.9779
ADCC-GJR-GARCH(1,1)	0.033	1.9779
CCC-GARCH(1,1)	0.033	1.9779
CCC-QGARCH(1,1)	0.037	1.2325
CCC-GJR-GARCH(1,1)	0.037	1.2325

Table 6.3: Results of the out-sample Kupiec test with critical value of the LR test statistic is 3.84 on the first 300 observations

Under 5% significant level, the multivariate models are still adequate in describing and forecasting the VaR. The different results in Table 6.2 and Table 6.3 suggest that when the multivariate conditional correlation models are used, the volatility forecast period should be appropriate to ensure accuracy and if it is permitted, one should update the parameters estimates once the new data comes available.

Chapter 7

Conclusion and Future Work

In the thesis, we studied multivariate conditional correlation GARCH models, including the CCC-GARCH, DCC-GARCH and ADCC-GARCH models. In constructing the conditional variances, three types of univariate GARCH models were employed: GARCH, the QGARCH and GJR-GARCH model. These models were applied to the daily returns of the General Electric Company and American Express Company.

In the parameter estimation of dynamic correlation models, we adopted the two steps method suggested by Engle [8]. Instead of the regular maximum likelihood estimation, we implemented a Markov Chain Monte Carlo algorithm, Metropolis within Gibbs algorithm.

After estimating the parameters of each model in the empirical application, we first studied the adequacy of the models and found that all of them are capable of describing the heteroskedasticity in the data. We also compared goodness of fit of the models by maximum value of the likelihood function and Bayesian Information Criterion. They gave different result. The likelihood function suggested the CCC-QGARCH(1,1) model while BIC favoured the CCC-GARCH(1,1) model. Finally, all the models were applied to

compute VaR and the Kupiec test was used to check the performance of the models. The in-sample test showed that the univariate models are inadequate in computing VaR. According to the out-of-sample test, all the models are able to forecast and estimate the VaR, but the out-of-sample period should be short for the conditional correlation GARCH models.

There are two points that can be improved in future work.

1. In the estimation of the DCC-GARCH model, we always estimated the parameter matrices $\bar{\mathbf{Q}}$ by $\frac{1}{n} \sum_{t=1}^T \boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t'$. An idea is whether the estimated correlation matrix \mathbf{R} from the CCC-GARCH model can be used instead. In a numerical experiment, the resulting two-step estimate performed better than the one suggested by Engle [8], but it lacks the theoretical support, which may be developed in future work.
2. Instead of the two-step method, one can estimate the parameters of the DCC-GARCH or ADCC-GARCH models by implementing a MCMC algorithm on the likelihood function as a whole to eliminate the error caused by two step methods. The Metropolis within Gibbs algorithm has been tried but the speed of convergence turned out to be slow, so that some other algorithm need to be employed.

Appendix A

A.1 Metropolis Hasting algorithm

Algorithm 1: Metropolis Hasting algorithm

input : An initial value $\boldsymbol{\theta}^{(0)}$ such that $\pi(\boldsymbol{\theta}^{(0)}) > 0$ and the number of iterations N .
Result: To have N enough target distribution samples to approximate the true value.

Initialization;

```
for  $i = 1 : N$  do
  current state  $\boldsymbol{\theta}^{cur}$ ;
  sample a posterior  $\boldsymbol{\theta}'$ ;
  sample random variable  $u$  from standard uniform distribution;
  if  $u < p = \min\{1, \frac{\pi(\boldsymbol{\theta}'|\mathbf{y}) \cdot q(\boldsymbol{\theta}^{cur}|\boldsymbol{\theta}')}{\pi(\boldsymbol{\theta}^{cur}|\mathbf{y}) \cdot q(\boldsymbol{\theta}'|\boldsymbol{\theta}^{cur})}\}$  then
     $\boldsymbol{\theta}^{cur} = \boldsymbol{\theta}'$ ;
     $\boldsymbol{\theta}^{(i)} = \boldsymbol{\theta}$ ;
  end
end
```

A.2 Gibbs sampling algorithm

Algorithm 2: Gibbs Sampling

input : An initial value $\boldsymbol{\theta}^{(0)}$ such that $\pi(\boldsymbol{\theta}^{(0)}) > 0$ and the number of iterations N .

Result: To have N enough target distribution samples to approximate the true value.

Initialization;

$\boldsymbol{\theta}^{cur} \leftarrow \boldsymbol{\theta}^{(0)}$;

for $i = 1 : N$ **do**

for $j = 1 : d$ **do**

 draw a new value for the j th component θ_j^{cur} of $\boldsymbol{\theta}^{cur}$ from the posterior full conditional $\pi_{\theta_j|\boldsymbol{\theta}_{-j}}(\theta_j|\boldsymbol{\theta}_{-j}^{cur})$

end

$\boldsymbol{\theta}^{i+1} \leftarrow \boldsymbol{\theta}^{cur}$;

end

A.3 Metropolis Within Gibbs sampling algorithm

Algorithm 3: Metropolis Within Gibbs sampling algorithm

input : An initial value $\boldsymbol{\theta}^{(0)}$ such that $\pi(\boldsymbol{\theta}^{(0)}) > 0$ and the number of iterations N .

Result: To have N enough target distribution samples to approximate the true value.

Initialization;

$\boldsymbol{\theta}^{cur} \leftarrow \boldsymbol{\theta}^{(0)}$;

for $i = 1 : N$ **do**

for $j = 1 : d$ **do**

 current state θ_j^{cur} ;

 sample a posterior θ'_j ;

 sample random variable u from standard uniform distribution;

if $u < p = \min\{1, \frac{\pi(\theta'_j|\mathbf{y}) \cdot q(\boldsymbol{\theta}^{cur}|\theta'_j)}{\pi(\theta_j^{cur}|\mathbf{y}) \cdot q(\theta'_j|\theta_j^{cur})}\}$ **then**

$\theta_j^{cur} = \theta'_j$;

$\theta_j^{(i)} = \theta_j^{cur}$;

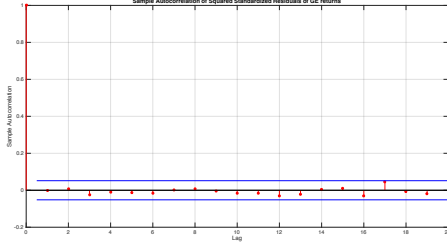
end

end

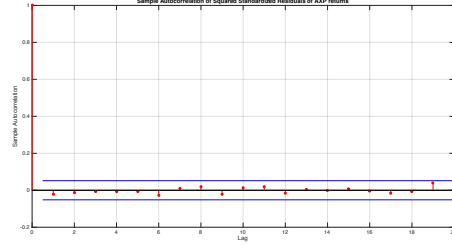
$\boldsymbol{\theta}^{i+1} \leftarrow \boldsymbol{\theta}^{cur}$;

end

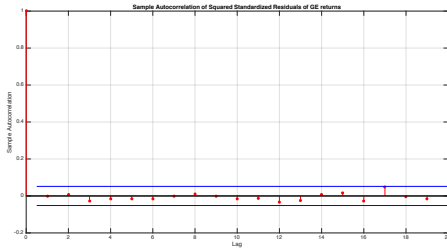
A.4 Sample ACF of standardised residuals



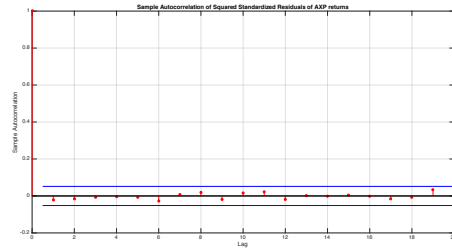
(a) Sample ACF of GE returns residuals standardised by univariate GARCH(1,1)



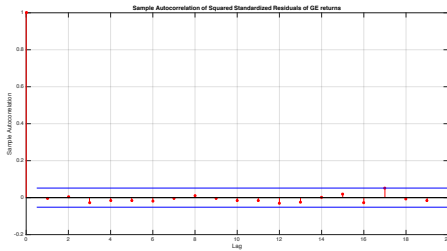
(b) Sample ACF of AXP returns residuals standardised by univariate GARCH(1,1)



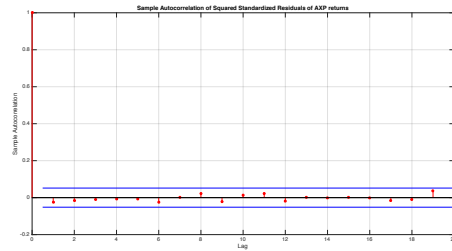
(a) Sample ACF of GE returns residuals standardised by univariate QGARCH(1,1)



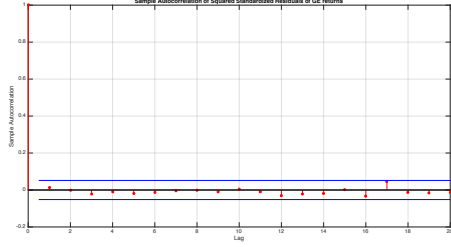
(b) Sample ACF of AXP returns residuals standardised by univariate QGARCH(1,1)



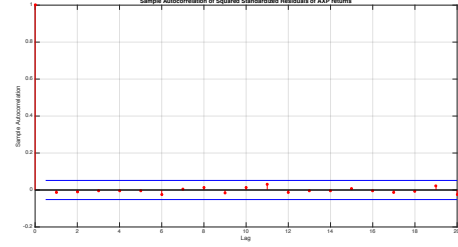
(a) Sample ACF of GE returns residuals standardised by univariate GJR-GARCH(1,1)



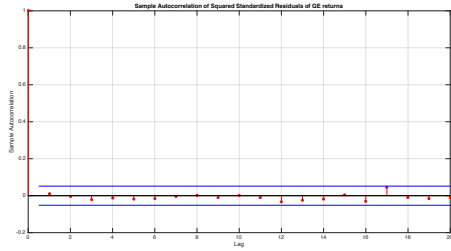
(b) Sample ACF of AXP returns residuals standardised by univariate GJR-GARCH(1,1)



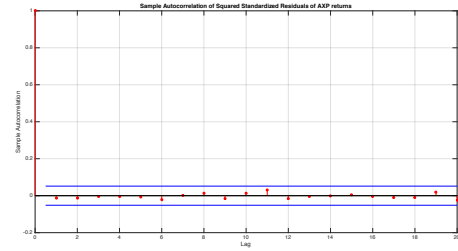
(a) Sample ACF of GE returns residuals standardised by CCC-GARCH(1,1)



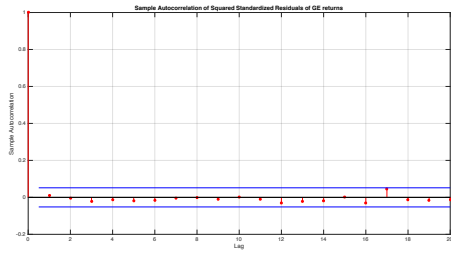
(b) Sample ACF of AXP returns residuals standardised by CCC-GARCH(1,1)



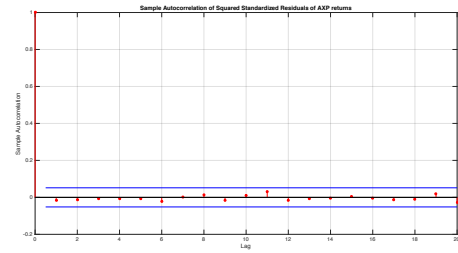
(a) Sample ACF of GE returns residuals standardised by CCC-QGARCH(1,1)



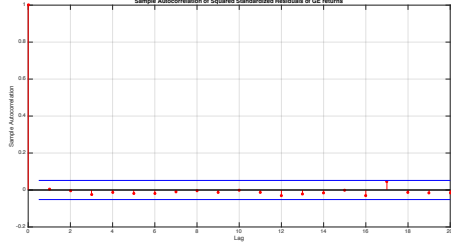
(b) Sample ACF of AXP returns residuals standardised by CCC-QGARCH(1,1)



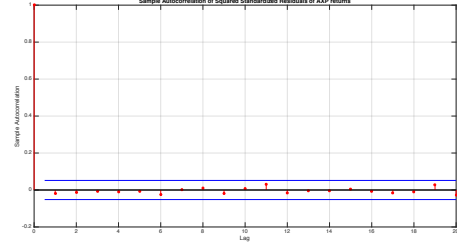
(a) Sample ACF of GE returns residuals standardised by CCC-GJR-GARCH(1,1)



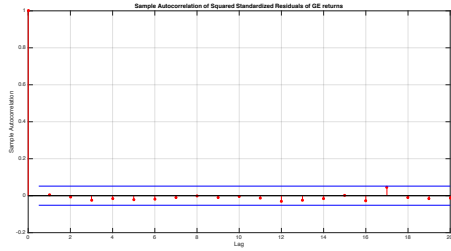
(b) Sample ACF of AXP returns residuals standardised by CCC-GJR-GARCH(1,1)



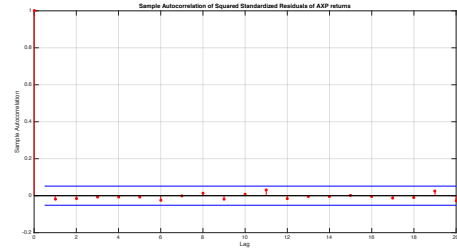
(a) Sample ACF of GE returns residuals standardised by DCC-GARCH(1,1)



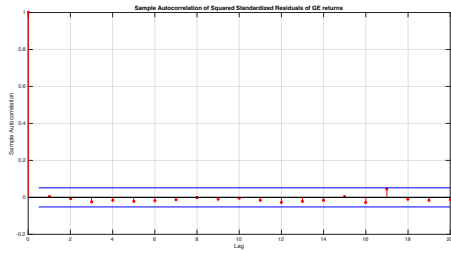
(b) Sample ACF of AXP returns residuals standardised by DCC-GARCH(1,1)



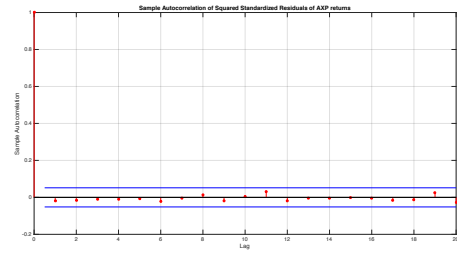
(a) Sample ACF of GE returns residuals standardised by DCC-QGARCH(1,1)



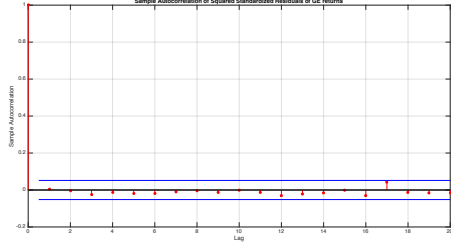
(b) Sample ACF of AXP returns residuals standardised by DCC-QGARCH(1,1)



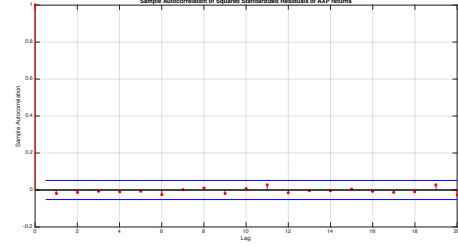
(a) Sample ACF of GE returns residuals standardised by DCC-GJR-GARCH(1,1)



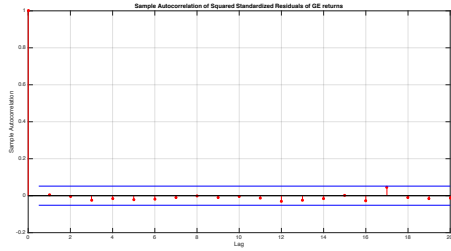
(b) Sample ACF of AXP returns residuals standardised by DCC-GJR-GARCH(1,1)



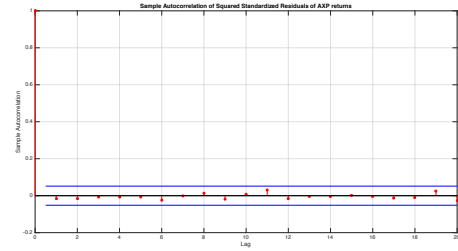
(a) Sample ACF of GE returns residuals standardised by ADCC-GARCH(1,1)



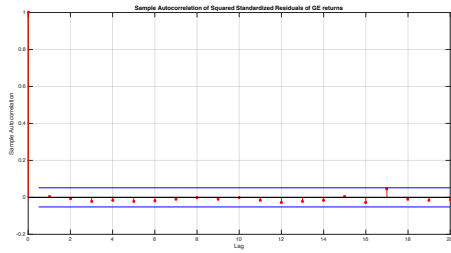
(b) Sample ACF of AXP returns residuals standardised by ADCC-GARCH(1,1)



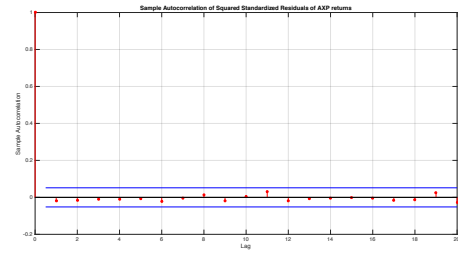
(a) Sample ACF of GE returns residuals standardised by ADCC-QGARCH(1,1)



(b) Sample ACF of AXP returns residuals standardised by ADCC-QGARCH(1,1)

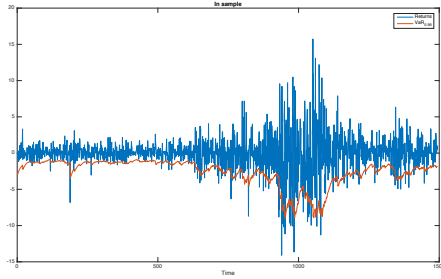


(a) Sample ACF of GE returns residuals standardised by ADCC-GJR-GARCH(1,1)

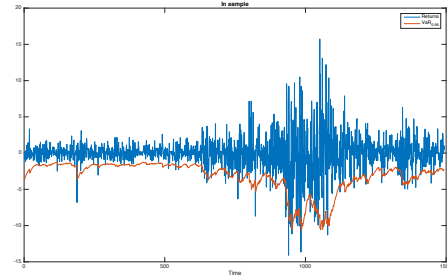


(b) Sample ACF of AXP returns residuals standardised by ADCC-GJR-GARCH(1,1)

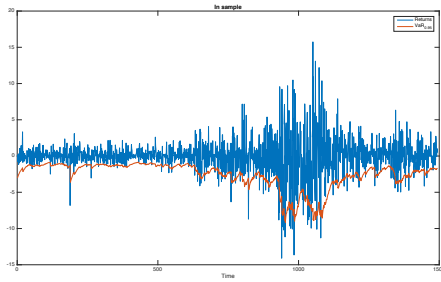
A.5 In-sample VaR computation



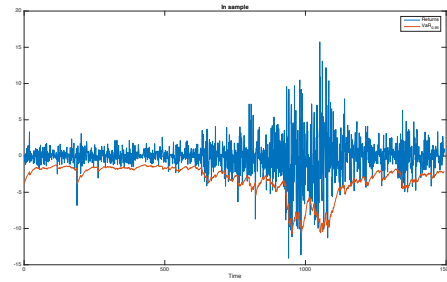
(a) In-sample $\text{VaR}_{0.95}$ by GARCH(1,1)



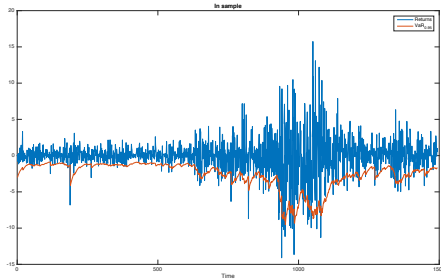
(b) In-sample $\text{VaR}_{0.95}$ by CCC-GARCH(1,1)



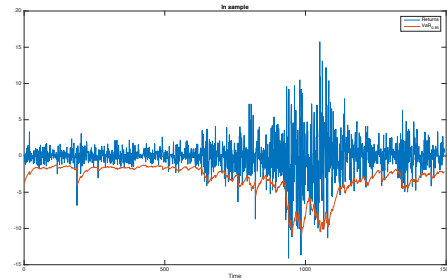
(a) In-sample $\text{VaR}_{0.95}$ by QGARCH(1,1)



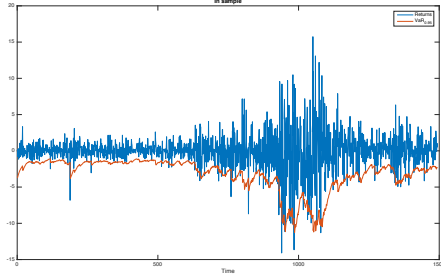
(b) In-sample $\text{VaR}_{0.95}$ by CCC-QGARCH(1,1)



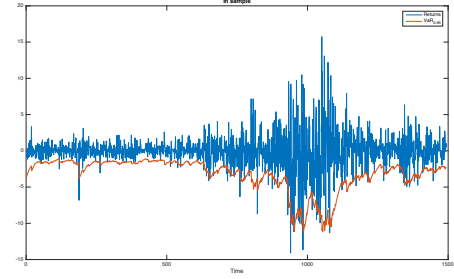
(a) In-sample $\text{VaR}_{0.95}$ by GJR-GARCH(1,1)



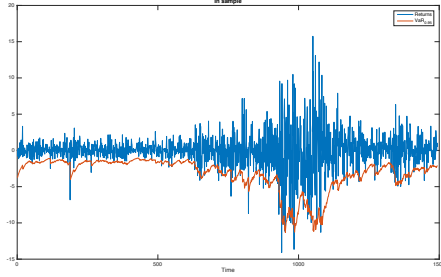
(b) In-sample $\text{VaR}_{0.95}$ by CCC-GJR-GARCH(1,1)



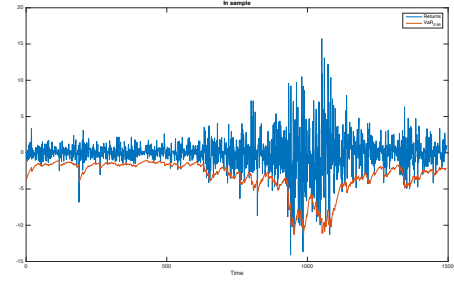
(a) In-sample $\text{VaR}_{0.95}$ by DCC-GARCH(1,1)



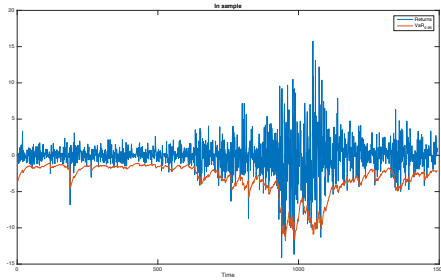
(b) In-sample $\text{VaR}_{0.95}$ by ADCC-GARCH(1,1)



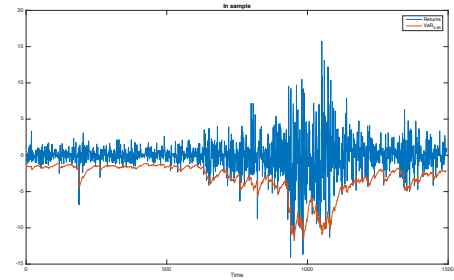
(a) In-sample $\text{VaR}_{0.95}$ by DCC-QGARCH(1,1)



(b) In-sample $\text{VaR}_{0.95}$ by ADCC-QGARCH(1,1)

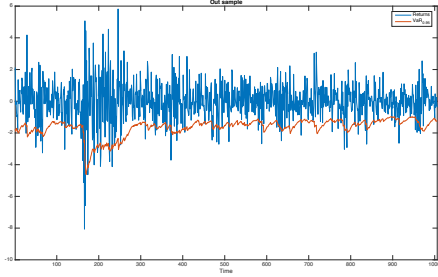


(a) In-sample $\text{VaR}_{0.95}$ by DCC-GJR-GARCH(1,1)

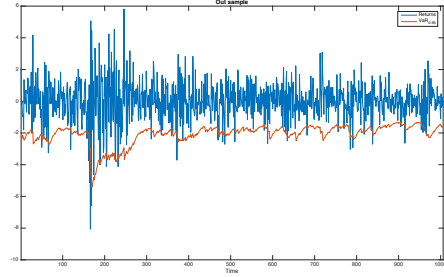


(b) In-sample $\text{VaR}_{0.95}$ by ADCC-GJR-GARCH(1,1)

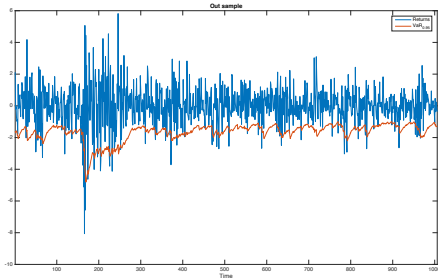
A.6 Out-sample VaR computation



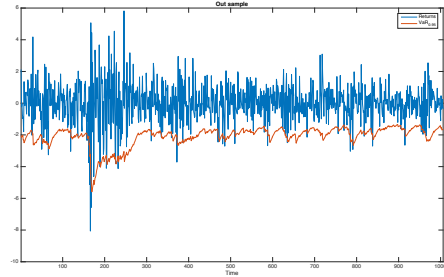
(a) Out-sample $\text{VaR}_{0.95}$ by GARCH(1,1)



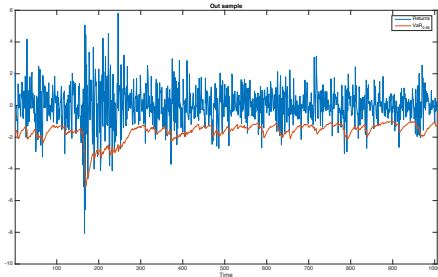
(b) Out-sample $\text{VaR}_{0.95}$ by CCC-GARCH(1,1)



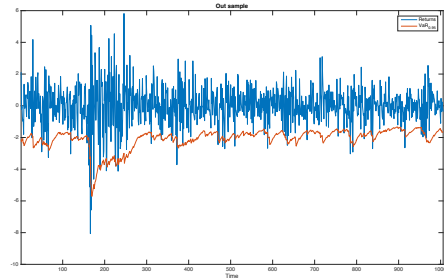
(a) Out-sample $\text{VaR}_{0.95}$ by QGARCH(1,1)



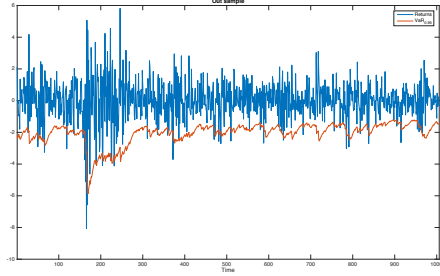
(b) Out-sample $\text{VaR}_{0.95}$ by CCC-QGARCH(1,1)



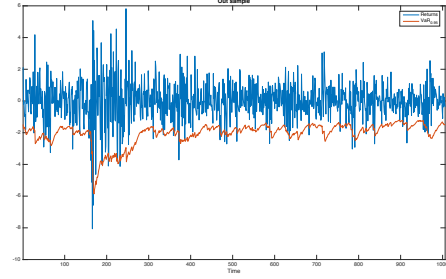
(a) Out-sample $\text{VaR}_{0.95}$ by GJR-GARCH(1,1)



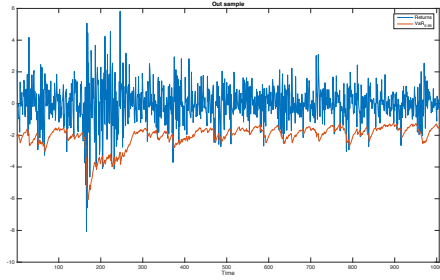
(b) Out-sample $\text{VaR}_{0.95}$ by CCC-GJR-GARCH(1,1)



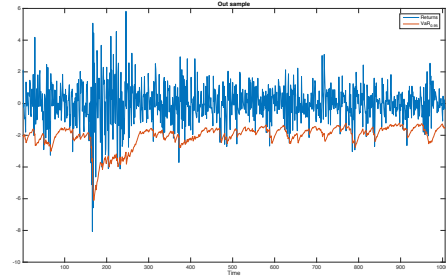
(a) Out-sample $\text{VaR}_{0.95}$ by DCC-GARCH(1,1)



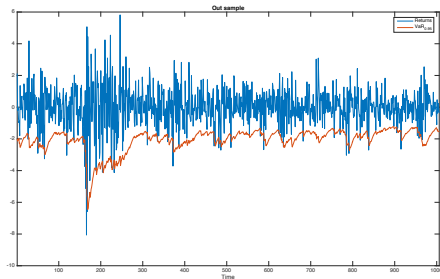
(b) Out-sample $\text{VaR}_{0.95}$ by ADCC-GARCH(1,1)



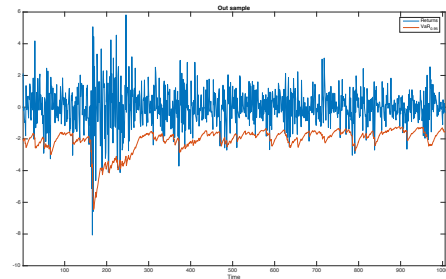
(a) Out-sample $\text{VaR}_{0.95}$ by DCC-QGARCH(1,1)



(b) Out-sample $\text{VaR}_{0.95}$ by ADCC-QGARCH(1,1)



(a) Out-sample $\text{VaR}_{0.95}$ by DCC-GJR-GARCH(1,1)



(b) Out-sample $\text{VaR}_{0.95}$ by ADCC-GJR-GARCH(1,1)

A.7 Mode estimates

GE&AXP Daily returns Jan 1st2005-December 9th 2010 (1,496 obs.)			
$\mathbf{y}_t = \mathbf{H}_t^{1/2} \mathbf{z}_t$			
$h_{11,t} = \omega_1 + \alpha_1 y_{1,t-1}^2 + \phi_1 y_{1,t-1} + \gamma_1 I[y_{1,t-1} < 0] + \beta_1 h_{11,t-1}$			
$h_{22,t} = \omega_2 + \alpha_2 y_{2,t-1}^2 + \phi_2 y_{2,t-1} + \gamma_2 I[y_{2,t-1} < 0] + \beta_2 h_{22,t-1}$			
$h_{ij,t} = 0, \quad \text{when } i \neq j$			
Parameter	GARCH(1,1)	QGARCH(1,1)	GJR-GARCH(1,1)
ω_1	0.0194	0.0263	0.0212
α_1	0.0889	0.0942	0.0726
ϕ_1		-0.0746	
γ_1			0.0356
β_1	0.9107	0.9043	0.9093
ω_2	0.0352	0.0482	0.0427
α_2	0.0960	0.1030	0.0548
ϕ_2		-0.1097	
γ_2			0.0923
β_2	0.9052	0.8973	0.9009

Table A.1: Parameter estimates based on modes of posterior samples for univariate GARCH(1,1) models.

GE&AXP Daily returns Jan 1st2005-December 9th 2010 (1,496 obs.)			
$\mathbf{y}_t = \mathbf{H}_t^{1/2} \mathbf{z}_t$			
$h_{11,t} = \omega_1 + \alpha_1 y_{1,t-1}^2 + \phi_1 y_{1,t-1} + \gamma_1 I[y_{1,t-1} < 0] + \beta_1 h_{11,t-1}$			
$h_{22,t} = \omega_2 + \alpha_2 y_{2,t-1}^2 + \phi_2 y_{2,t-1} + \gamma_2 I[y_{2,t-1} < 0] + \beta_2 h_{22,t-1}$			
$h_{ij,t} = \rho, \quad \text{when } i \neq j$			
Parameter	GARCH(1,1)	QGARCH(1,1)	GJR-GARCH(1,1)
ω_1	0.0231	0.0280	0.0246
α_1	0.0715	0.0756	0.0703
ϕ_1		-0.0483	
γ_1			0.0096
β_1	0.9227	0.9175	0.9193
ω_2	0.0522	0.0577	0.0549
α_2	0.0819	0.0823	0.0632
ϕ_2		-0.0853	
γ_2			0.0404
β_2	0.9108	0.9098	0.9094
ρ	0.5857	0.5805	0.5842

Table A.2: Parameter estimates based on modes of posterior samples for CCC-GARCH models.

GE&AXP Daily returns Jan 1st 2005-December 9th 2010 (1,496 obs.)			
$Q_t = (1 - a - b)\bar{Q} + a\epsilon_{t-1}\epsilon'_{t-1} + bQ_{t-1}$			
Parameter	GARCH(1,1)	QGARCH(1,1)	GJR-GARCH(1,1)
a	0.0295	0.0279	0.0281
b	0.8113	0.8218	0.7776
\bar{Q}	$\begin{bmatrix} 0.9801 & 0.5669 \\ 0.5669 & 0.9872 \end{bmatrix}$	$\begin{bmatrix} 0.9800 & 0.5614 \\ 0.5614 & 0.9852 \end{bmatrix}$	$\begin{bmatrix} 0.9786 & 0.5596 \\ 0.5596 & 0.9839 \end{bmatrix}$

Table A.3: Parameter estimates based on modes of posterior samples for DCC-GARCH models based on the mode of univariate models in Table A.1

GE&AXP Daily returns Jan 1st 2005-December 9th 2010 (1,496 obs.)			
$Q_t = (\bar{Q} - a\bar{Q} - b\bar{Q} - g\bar{N}) + a\epsilon_{t-1}\epsilon'_{t-1} + bQ_{t-1} + g\mathbf{n}_{t-1}\mathbf{n}_{t-1}$			
Parameter	GARCH(1,1)	QGARCH(1,1)	GJR-GARCH(1,1)
a	0.0249	0.0227	0.0230
b	0.7914	0.8362	0.8359
g	0.0196	0.0221	0.0242
\hat{Q}	$\begin{bmatrix} 0.9785 & 0.5661 \\ 0.5661 & 0.9859 \end{bmatrix}$	$\begin{bmatrix} 0.9782 & 0.5614 \\ 0.5614 & 0.9848 \end{bmatrix}$	$\begin{bmatrix} 0.9781 & 0.5586 \\ 0.5586 & 0.9822 \end{bmatrix}$
\hat{N}	$\begin{bmatrix} 0.4848 & 0.3192 \\ 0.3192 & 0.5184 \end{bmatrix}$	$\begin{bmatrix} 0.4864 & 0.3192 \\ 0.3192 & 0.5184 \end{bmatrix}$	$\begin{bmatrix} 0.4663 & 0.1747 \\ 0.1747 & 0.5418 \end{bmatrix}$

Table A.4: Parameter estimates based on modes of posterior samples for ADCC-GARCH models based on the mode of univariate models in Table A.1

References

- [1] David Ardia. *Financial risk management with Bayesian estimation of GARCH models*. Springer, 2008.
- [2] David Ardia and Lennart F Hoogerheide. Bayesian estimation of the garch (1, 1) model with student-t innovations. *The R Journal*, 2(2):41–47, 2010.
- [3] Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327, 1986.
- [4] Tim Bollerslev. Modelling the coherence in short-run nominal exchange rates: a multivariate generalized arch model. *The Review of Economics and Statistics*, pages 498–505, 1990.
- [5] Tim Bollerslev, Robert F Engle, and Jeffrey M Wooldridge. A capital asset pricing model with time-varying covariances. *The Journal of Political Economy*, pages 116–131, 1988.
- [6] Lorenzo Cappiello, Robert F Engle, and Kevin Sheppard. Asymmetric dynamics in the correlations of global equity and bond returns. *Journal of Financial econometrics*, 4(4):537–572, 2006.

- [7] Cousineau Denis and Helie Sebastien. Improving maximum likelihood estimation using prior probabilities: A tutorial on maximum a posteriori estimation and an examination of the weibull distribution. *Tutorials in Quantitative Methods for Psychology*, 9(2):61–71, 2013.
- [8] Robert Engle. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3):339–350, 2002.
- [9] Robert F Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007, 1982.
- [10] Robert F Engle and Kenneth F Kroner. Multivariate simultaneous generalized arch. *Econometric theory*, 11(01):122–150, 1995.
- [11] Robert F Engle and Victor K Ng. Measuring and testing the impact of news on volatility. *The journal of finance*, 48(5):1749–1778, 1993.
- [12] Stuart Geman and Donald Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (6):721–741, 1984.
- [13] Lawrence R Glosten, Ravi Jagannathan, and David E Runkle. On the relation between the expected value and the volatility of the nominal excess return on stocks. *The journal of finance*, 48(5):1779–1801, 1993.
- [14] W Keith Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.
- [15] Paul H Kupiec. Techniques for verifying the accuracy of risk measurement models. *THE J. OF DERIVATIVES*, 3(2), 1995.

- [16] Shiqing Ling and Michael McAleer. Stationarity and the existence of moments of a family of garch processes. *Journal of Econometrics*, 106(1):109–117, 2002.
- [17] Nicholas Metropolis, Arianna W Rosenbluth, Marshall N Rosenbluth, Augusta H Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092, 1953.
- [18] Daniel B Nelson. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, pages 347–370, 1991.
- [19] Gideon Schwarz et al. Estimating the dimension of a model. *The annals of statistics*, 6(2):461–464, 1978.
- [20] Enrique Sentana. Quadratic arch models. *The Review of Economic Studies*, 62(4):639–661, 1995.
- [21] Tetsuya Takaishi. Bayesian estimation of garch model by hybrid monte carlo. *arXiv preprint physics/0702240*, 2007.
- [22] Luke Tierney. Markov chains for exploring posterior distributions. *the Annals of Statistics*, pages 1701–1728, 1994.
- [23] Ruey S Tsay. *Analysis of financial time series*, volume 543. John Wiley & Sons, 2005.
- [24] Jean-Michel Zakoian. Threshold heteroskedastic models. *Journal of Economic Dynamics and control*, 18(5):931–955, 1994.
- [25] Eric Zivot. Practical issues in the analysis of univariate garch models. In *Handbook of financial time series*, pages 113–155. Springer, 2009.